ÜBER DAS RIVLIN - PROBLEM DER SIMULTANEN INVERSEN TSCHEBYSCHEFF-APPROXIMATION

Heinz-Gerhard Hegering

München
1971

ÜBER DAS RIVLIN - PROBLEM DER SIMULTANEN INVERSEN TSCHEBYSCHEFF-APPROXIMATION

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der Naturwissenschaftlichen Fakultät
der Ludwig-Maximilians-Universität
zu München

vorgelegt von

Heinz-Gerhard Hegering
aus Recklinghausen

München, im April 1971

UNIV.
LEIDEN
BIBL.

1. Berichterstatter: Prof. Dr. G. Hämmerlin
2. Berichterstatter: Prof. Dr. K. Jörgens

Tag der mündlichen Prüfung: 7.7.1971

Fotodruck: Mikrokopie GmbH, 8 München 22, Bruderstraße 9

INHALT

		Seite
Einleitung und Übersicht		1
§1	Das RIVLIN-Problem in normierten linearen Räumen; einige äquivalente Problemstellungen und allgemeine Charakterisierungen	8
§2	Erweiterte Tschebyscheff-Systeme; - Definitionen, Eigenschaften und Beispiele -	20
§3	Notwendige Kriterien für die Lösbarkeit bestimmter RIVLIN-Probleme im Raum C[a,b]	26
§4	Das RIVLIN-Problem im Raum C[a,b]- einige Charakterisierungssätze	37
§5	Beispiele	58
Literaturverzeichnis		72

EINLEITUNG und ÜBERSICHT

Die Aufgabe, gegebene komplizierte funktionale Zusammenhänge möglichst gut durch einfachere Funktionen bestimmter Bauart darzustellen, war der Ausgangspunkt für die klassische Problemstellung der Approximationstheorie, die sich - gleich auf allgemeine normierte lineare Räume übertragen - wie folgt beschreiben läßt:

Problem (P1): R sei ein normierter linearer Raum. G sei eine Teilmenge von R. $f \in R$ sei vorgegeben. Gesucht wird ein Element $\tilde{g} \in G$ mit der Eigenschaft $\|f-\tilde{g}\| \leq \|f-g\|$ für alle $g \in G$, also ein Element $\tilde{g} \in G$ mit

$$(\dagger) \qquad \|f-\tilde{g}\| = \inf_{g \in G} \|f-g\| =: E_G(f)$$

Diese allgemeine Formulierung von (P1) impliziert eine Reihe von Teilproblemen, so etwa

(a) Für welche Teilmengen G ist die Aufgabe (P1) lösbar? Das ist die Frage der Existenz von Elementen bester Approximation und der Charakterisierung von sogenannten "Existenzmengen" G.

(b) Falls Elemente $\tilde{g} \in G$ mit der Eigenschaft (†) existieren, sogenannte Minimallösungen für f aus G, sind sie eindeutig bestimmt?

(c) Wie sind Minimallösungen zu charakterisieren?

(d) Wie lassen sich Minimallösungen praktisch berechnen? Kann man allgemeine Konstruktionsverfahren angeben?

(e) Welche Aussagen lassen sich bezüglich der Minimalabweichung $E_G(f)$ und des Approximationsoperators treffen, also der Zuordnung, die jedem $f \in R$ die Elemente bester Approximation aus G zuordnet?

(f) Welche Antworten für die Fragen (a)-(e) ergeben sich bei speziellen Räumen R und speziellen Teilmengen $G \subset R$?

Das Problem (P1) ist ausführlich untersucht worden. Im Fall, daß G ein linearer Teilraum von R ist (→ lineare Approximationsaufgabe), sei auf die Monographie von I.SINGER [1970] und die dort zitierte Literatur hingewiesen, im Falle von nichtlinearen Teilmengen G (→ nichtlineare Approximation) z.B. auf die Arbeiten von B.BROSOWSKI [1968,1969] und K.H. HOFFMANN [1968].

Ausgehend vom Problem (P1), das anfänglich für den Raum $R=C[a,b]$ betrachtet worden war, ergaben sich jedoch schon zur Zeit der letzten Jahrhundertwende noch eine Reihe weiterer Aspekte für approximationstheoretische Untersuchungen, die inzwischen große Bedeutung erlangt haben, so das folgende

Problem (P2): Sei $R=C[a,b]$ der mit der Tschebyscheff-Norm versehene lineare Raum der auf dem endlichen Intervall $[a,b]$ definierten stetigen reellwertigen Funktionen. Gegeben seien die linearen Teilräume

$$G_n = P_n := \{p_n; p_n = \sum_{\ell=0}^{n} a_\ell x^\ell,\ a_\ell \in \mathbb{R}\} \subset R, n \in \mathbb{N}_0$$

bzw.

$$G_n = T_n^* := \{t_n; t_n = a_0 + \sum_{\ell=1}^{n} (a_\ell \sin \ell x + b_\ell \cos \ell x),$$

$$a_\ell \in \mathbb{R},\ b_\ell \in \mathbb{R}\} \subset R,\ n \in \mathbb{N} \text{ und } G_0 = T_0^* := \{a_0, a_0 \in \mathbb{R}\}.$$

Ferner sei ein Element $f \in R$ gegeben.

Bestimme die Elemente bester Approximation für f aus den Mengen G_n, (n=0,1,2,...), und charakterisiere das Verhalten der Folge $E_{G_n}(f)$. Hängt dieses Verhalten von Struktureigenschaften von f ab?

Der Satz von WEIERSTRASS aus dem Jahre 1885, der beweist, daß $\lim_{\ell\to\infty} E_{G_\ell}(f) = 0$ gilt, besagt, daß die Problemstellung (P2) sinnvoll ist. Antworten auf das Problem (P2) geben u.a. ferner die bekannten Sätze von D.JACKSON aus dem Jahre 1911 - etwa in D.JACKSON [1930] - (die von differentiellen Eigenschaften von f auf die Schnelligkeit schließen, mit der die Folge $E_{G_\ell}(f)$ gegen Null für $\ell\to\infty$ geht), und deren Erweiterungen von A.ZYGMUND [1945].
Die Antworten zu (P2) werden mitunter auch "direkte Sätze" genannt, im Unterschied zu gewissen Sätzen bezüglich des folgenden Problems (P3), die auch "indirekte Sätze" heißen, da (P3) in bestimmter Weise eine Umkehrung zu (P2) darstellt (vergleiche auch weiter unten).

Problem (P3): Sei $R = C[a,b]$ und G_n, $n \geq 0$, wie in (P2). Ferner sei eine monotone Nullfolge $\{\varepsilon_\ell,\ \ell \in \mathbb{N}_0\}$ gegeben.

(α) Existiert eine $f \in R$ mit $E_{G_\ell}(f) = \varepsilon_\ell$, $\ell \geq 0$?

(β) Lassen sich aus der Konvergenzgeschwindigkeit der Folge $\{\varepsilon_\ell\}$ Rückschlüsse ziehen auf Struktureigenschaften (Stetigkeit, Differenzierbarkeit etc.) eines solchen f?

S.N.BERNSTEIN [1938] bejahte die Frage (α) von (P3) und gab auch auf den Teil (β) eine die Frage bestätigende Antwort u.a. in seiner bekannten Arbeit S.N.BERNSTEIN [1912]. Erweiterungen dieser Fragestellung untersuchten A.ZYGMUND [1945], S.B.STEČKIN [1951] und M.ZAMANSKY [1949]. P.L.BUTZER und K.SCHERER [1969] zeigten unter gewissen Voraussetzungen die Äquivalenz der Aussagen von JACKSON, BERNSTEIN, STEČKIN und ZAMANSKY; ferner übertrugen sie die Ergebnisse, die die Zusammenhänge zwischen (P2) und (P3) aufweisen, auf allgemeine Banachräume. J.NITSCHE [1968] schließlich entwickelte Aussagen vom Bernstein- bzw. Jackson-Typ für projektive Verfahren, insbesondere für RITZ-Verfahren zur Lösung von Operatorgleichungen.

Wie schon angedeutet, ist (P3) eine zu (P2) inverse Fragestellung, insofern nämlich, daß nicht die zu approximierende Funktion $f \in R$ vorgegeben wird und dazu die Elemente bester Approximation aus gegebenen Teilmengen G_ℓ gesucht werden - das entspricht (P2) - , sondern Zahlenwerte vorgegeben werden, die die Minimalabstände einer zu suchenden Funktion $f \in R$ bezüglich der Mengen G_ℓ darstellen sollen.

(P2) läßt jedoch noch eine andere Umkehrung zu. Es war T.J.RIVLIN [1969], der 1968 anläßlich einer Tagung über Approximationstheorie in Oberwolfach die folgende weitere Möglichkeit der inversen Fragestellung der Situation von (P2) zur Diskussion stellte:

Problem (P4): Charakterisiere solche n-Tupel von algebraischen Polynomen $\{p_0, p_1, \ldots, p_{n-1}\}$ mit grad $p_\ell = \ell$, $\ell = 0(1)n-1$, für die ein $f \in C[-1,1]$ existiert, so daß das Polynom bester Approximation (im Sinne von Tschebyscheff) vom Grad ℓ bezüglich f gerade p_ℓ, $\ell = 0(1)n-1$, ist. Wie lautet die Antwort im Spezialfall n=2?

Dieses Problem bedeutet also das Auffinden eines Elementes $f \in C[-1,1]$ bei mehreren "vorgeschriebenen Minimallösungen" p_ℓ, die zudem noch aus verschiedenen Teilmengen von $C[-1,1]$ stammen. Es liegt somit eine "simultane inverse Aufgabenstellung" der Approximationstheorie vor.

Dieser Problematik wollen wir uns in der vorliegenden Arbeit widmen.

Dabei ergeben sich unter anderem folgende Teilprobleme:

(1) Welchen Forderungen müssen die Polynome p_ℓ genügen, damit ein solches gesuchtes Element f existieren kann? Das ist die Frage nach notwendigen Kriterien. Gibt es hinreichende Kriterien für die Existenz eines solchen f, das wir im Folgenden auch RIVLIN-Funktion nennen werden? Lassen sich in Spezialfällen Charakterisierungen finden?

(2) Ist es möglich, ein eventuell im Sinne von (P4) existierendes Element f konstruktiv anzugeben? Ist eine RIVLIN-Funktion eindeutig bestimmt?

(3) Beeinflussen in bestimmten Fällen Stetigkeits- und Differenzierbarkeitseigenschaften der Vorgabepolynome die entsprechenden Eigenschaften der RIVLIN-Funktion?

(4) Welche Ergebnisse bezüglich der Fragen (1)-(3) ergeben sich, falls man im Problem (P4) nicht nur Polynome als "Vorgabefunktionen" zuläßt, sondern z.B. Funktionen aus den allgemeinen Haar-Unterräumen?

(5) Lassen sich Aussagen zu Frage (1) machen, falls man das von T.J.RIVLIN gestellte Ausgangsproblem (P4) entsprechend auf allgemeine normierte lineare Räume überträgt?

T.J.RIVLIN selbst hat bezüglich der Frage (1) ein notwendiges Kriterium angegeben. Weiter haben sich F.DEUTSCH, P.MORRIS, I.SINGER [1969] und D.SPRECHER [1969,1969a] mit der Aufgabenstellung von RIVLIN befaßt und u.a. das Problem (P4) für den Spezialfall n = 2 gelöst. Von D.SPRECHER stammt ein weiteres notwendiges Kriterium für eine RIVLIN-Funktion und für algebraische Polynome je eine Charakterisierung im Falle eines beliebigen vorgegebenen Polynompaares (p_i, p_ℓ) und im Falle des speziellen gewöhnlichen Polynomtripels (p_0, p_1, p_2).
DEUTSCH, MORRIS, SINGER [1969] übertrugen die Fragestellung (P4) auf den Fall eines beliebigen normierten linearen Raumes mit Approximationen bezüglich gewisser linearer Teilräume. Unter anderem bewiesen sie dabei einen Charakterisierungssatz, der in der vorliegenden Arbeit des öfteren benötigt und der deshalb an späterer Stelle (→ Satz 1.5) dargestellt wird. Im übrigen jedoch werden in der Arbeit von DEUTSCH, MORRIS, SINGER [1969] im Wesentlichen noch eine Reihe von Ergebnissen entwickelt unter der Voraussetzung, daß der Approximationsoperator linear ist.
Ergebnisse mit dieser Voraussetzung sind aber nicht anwendbar auf das RIVLIN-Problem im Raum $C[a,b]$, da die Existenz

von linearen Approximationsoperatoren gerade charakteristisch ist für das Vorliegen von Hilberträumen, wie J.STOER [1967] gezeigt hat. (Der Aspekt der Anwendbarkeit der Ergebnisse auf den mit der Tschebyscheffnorm versehenen Raum C[a,b] soll aber auch im Folgenden im Auge behalten werden).

Die hier vorliegende Arbeit beschäftigt sich mit der Fragestellung von T.J.RIVLIN in verallgemeinerter Form. Dabei lassen sich die Inhalte der einzelnen Paragraphen wie folgt kurz umreißen:

Im ersten Paragraphen wird zunächst das RIVLIN-Problem (P4) übertragen auf allgemeine normierte lineare Räume. Mithilfe bekannter Charakterisierungssätze aus der Approximationstheorie, die Antwort auf die Frage (c) zu (P1) geben, werden Sätze bezüglich des verallgemeinerten RIVLIN-Problems (P) gewonnen, die einerseits (implizit) Antworten auf die obige Fragestellung (5) geben, andererseits aber auch als zu (P) äquivalente Problemformulierungen aufgefaßt werden können, und die sich so als Schlüssel zu weiteren (expliziten) Charakterisierungen anbieten. Diese Möglichkeit wird im Verlaufe der Arbeit, insbesondere in den Paragraphen 3 und 4, genutzt.
Nachdem im zweiten Paragraphen (erweiterte) Tschebyscheff-Systeme definiert, einige ihrer Eigenschaften beschrieben und Beispiele für diese Funktionensysteme angegeben werden, sind die Hilfsmittel bereitgestellt für die folgenden Abschnitte.
Der dritte und vierte Paragraph beschäftigt sich mit der Fragestellung von T.J.RIVLIN im Raum C[a,b] versehen mit der Tschebyscheff-Norm. Das "simultane inverse Problem" wird jedoch bezüglich allgemeinerer Teilräume als denen der algebraischen Polynome, nämlich bei Haarsystemen und erweiterten Tschebyscheff-Systemen untersucht. Dabei finden Ergebnisse aus §1 Anwendung. In §3 interessieren wir uns für notwendige Bedingungen, also für Kriterien, die erfüllt sein

müssen, falls eine RIVLIN-Funktion bezüglich gewisser vorgegebener Funktionen (sogenannter u-Polynome oder verallgemeinerter Polynome von (E)T-Systemen) existiert. Als Ergebnis erhalten wir neue notwendige Bedingungsgruppen für die in §2 eingeführten Funktionenklassen. Auch lassen sich die weiter oben erwähnten notwendigen Kriterien von D.SPRECHER und T.J.RIVLIN auf diese Situation übertragen.
Im vierten Paragraphen werden neue Charakterisierungssätze bewiesen, die die erwähnten Charakterisierungen von D.SPRECHER [1969] für gewöhnliche Polynompaare und für das gewöhnliche spezielle Polynomtripel (p_0,p_1,p_2), vergleiche D.SPRECHER [1969a], als Spezialfälle enthalten. Die Existenzbeweise für eine RIVLIN-Funktion werden dabei durchweg konstruktiv geführt.
Insgesamt beinhalten die Paragraphen 1,3 und 4 jeweils Teilantworten auf die Detailfragen (1)-(5), die weiter oben angedeutet wurden.

Der abschließende fünfte Paragraph untersucht die Frage, inwieweit bestimmte vorher entwickelte Kriterien für die Existenz von RIVLIN-Funktionen $f \in C[a,b]$ erfüllt sind, wenn als "vorgeschriebene Minimallösungen" die Repräsentanten einiger bekannter Funktionsklassen (Monome, Tschebyscheff-Polynome 1. und 2. Art, Legendre-Polynome, sin-cos-Terme) gewählt werden, die u.a. in der Approximationstheorie häufig Verwendung finden. Es zeigt sich, daß dabei die Tschebyscheff-Polynome zweiter Art hinsichtlich dieser Fragestellung eine besondere Rolle spielen.

An dieser Stelle möchte ich mich herzlich bei Herrn Professor Dr. G. Hämmerlin bedanken für wertvolle Hinweise und sein Interesse am Fortgang dieser Arbeit.
Ebenso danke ich Herrn Dr. K.-H.Hoffmann für Anregungen und klärende Diskussionen.

§1 Das RIVLIN-Problem in normierten linearen Räumen

- Einige äquivalente Problemstellungen und allgemeine Charakterisierungen -

Die Fragestellung von T.J.RIVLIN [1969] betrachtete Tschebyscheff-Approximationen im Raum $C[a,b]$ der stetigen und reellwertigen, auf dem endlichen Intervall $[a,b]$ definierten Funktionen mittels algebraischer Polynome. Verallgemeinernd wird im vorliegenden Paragraphen dieses Problem in allgemeinen normierten linearen Räumen behandelt.

Dazu wird zunächst die Ausgangs-Fragestellung von T.J.RIVLIN auf diese Situation übertragen und dann versucht, einige der bekanntesten Sätze der Approximationstheorie, die eine Charakterisierung von Elementen bester Approximation (Minimallösungen) betreffen, anzuwenden auf das verallgemeinerte RIVLIN-Problem.

Man erhält so allgemeine Charakterisierungssätze für das verallgemeinerte RIVLIN-Problem, die jedoch zugleich als zur RIVLIN-Fragestellung äquivalente Problemformulierungen interpretiert werden können, und die so einen Ausgangspunkt für weitere Untersuchungen darstellen.

Bezeichne also im Folgenden R einen normierten linearen Raum und R^* seinen topologischen Dualraum, normiert durch $\|f^*\| = \sup_{f\in R, \|f\|\leq 1} |f^*(f)|$.

Sei $\{G_i,\ i\in I\subset\mathbb{N}\}$ eine Familie von Teilmengen von R. Für $f\in R$ sei der Minimalabstand bezüglich G_i $(i\in I)$ definiert durch

$$E_{G_i}(f) := \inf_{g\in G_i} \|f-g\|.$$

Ein Element $\tilde{g}\in G_i$ mit der Eigenschaft $\|\tilde{g}-f\| = E_{G_i}(f)$ heißt Minimallösung für f aus G_i oder auch Element bester Approximation aus G_i an f. Die Menge der Minimallösungen für ein $f\in R$ bzgl. einer Teilmenge G_i ist die Menge

$$\mathcal{P}_{G_i}(f) := \{\tilde{g} \in G_i \text{ mit } \|\tilde{g}-f\| = E_{G_i}(f)\}$$

Für Elemente $f \in R$ mit der Eigenschaft $\mathcal{P}_{G_i}(f) \neq \phi$ ist der Approximationsoperator $A_{G_i} : R \to G_i$ definiert durch die Vorschrift $A_{G_i}(f) = \mathcal{P}_{G_i}(f) \subset G_i$; i.A. handelt es sich dabei um eine mengenwertige Abbildung. Gibt es Punkte $f \in R$, die kein Element bester Approximation aus G_i besitzen, d.h. ist also G_i nicht proximal, so schreibt man $A_{G_i}(f) = \phi$ für solche $f \in R$.
Für ein festes Element $g_i \in G_i$ $(i \in I)$ sei weiter definiert

$$L(g_i) := \{f \in R \text{ mit } E_{G_i}(f) = \|f-g_i\|\} = \{f \in R \text{ mit } g_i \in A_{G_i}(f)\}.$$

Mit diesen Bezeichnungen und Definitionen lautet nun das
verallgemeinerte RIVLIN-Problem (P)

(P): In R sei eine Familie $\{G_i, i \in I \subset \mathbb{N}\}$ von Teilmengen aus R vorgegeben. Charakterisiere Auswahlen $\{g_i, i \in I\}$ von Elementen $g_i \in G_i$, $i \in I$, mit der Eigenschaft $\bigcap_{i \in I} L(g_i) \neq \phi$.
Charakterisiere Elemente $f \in R$ mit $f \in \bigcap_{i \in I} L(g_i)$.

Definition 1.1: Es liege die RIVLIN-Aufgabe (P) vor. Ein Element $f \in R$ mit $f \in \bigcap_{i \in I} L(g_i)$ nennen wir RIVLIN-Funktion (bezüglich der Folge $\{g_i, i \in I\}$). Im Falle einer endlichen Indexmenge $I = \{1,\ldots,n\}$ schreiben wir auch $L(g_1,\ldots,g_n)$ statt $\bigcap_{i=1}^{n} L(g_i)$.

Die nun folgenden Sätze 1.1, 1.2, 1.4, 1.5 betrachten das Problem (P) für den Fall, daß die vorgegebenen Teilmengen G_i, $i \in I$, lineare Teilräume von R sind. Satz 1.3 läßt für das Problem (P) umfassendere Teilmengen zu.

Satz 1.1 charakterisiert eine RIVLIN-Funktion f durch gewisse lineare stetige Funktionale $f_i^* \in R^*$, die jeweils auf dem Teilraum G_i verschwinden und Abweichungsfunktionale des Elementes $f-g_i$ sind:

SATZ 1.1: R sei ein normierter linearer Raum mit dem topologischen Dualraum R^*. G_i, $i \in I \subset \mathbb{N}$, seien lineare Teilräume von R mit $\overline{G}_i \neq R$. Seien Elemente $g_i \in G_i$, $i \in I$, vorgegeben und $f \in R$. Dann gilt: Es ist $f \in \bigcap_{i\in I} L(g_i)$ genau dann, wenn für jedes $i \in I$ ein Funktional $f_i^* \in R^*$ existiert mit folgenden Eigenschaften

$$(1.1) \quad \|f_i^*\| = 1$$

$$(1.2) \quad f_i^*(g) = 0 \text{ für alle } g \in G_i$$

$$(1.3) \quad f_i^*(f-g_i) = \|f-g_i\|$$

Beweis: Hinlänglichkeit:

Existiere für jedes $i \in I$ ein Funktional $f_i^* \in R^*$, welches den Bedingungen (1.1),(1.2) und (1.3) genügt. Mit $g \in G_{i_o}$, $i_o \in I$ beliebig, gilt dann

$$\|f-g_{i_o}\| \overset{(3)}{=} |f_{i_o}^*(f-g_{i_o})| \overset{(2)}{=} |f_{i_o}^*(f-g)| \leq \|f_{i_o}^*\| \overset{(1)}{\|f-g\|} = \|f-g\|,$$

somit ist $g_{i_o} \in \mathcal{P}_{G_{i_o}}(f)$, also $f \in L(g_{i_o})$. Da i_o beliebig aus I gewählt war, gilt damit $f \in L(g_i)$ für alle $i \in I$, d.h. es ist $f \in \bigcap_{i\in I} L(g_i)$.

Notwendigkeit der Bedingungen:

Sei $f \in \bigcap_{i\in I} L(g_i)$, dann ist für alle $i \in I$: $f \in L(g_i)$, d.h. $g_i \in \mathcal{P}_{G_i}(f)$, und somit

$$\|f-g_i\| = d(f,G_i) := \inf_{g\in G_i} \|f-g\|.$$

Da für jeden linearen Teilraum $G \subset R$ und jedes Element $f \in R$ gilt

$$(1.4) \quad \mathcal{P}_G(f) = \begin{cases} f & \text{für } f \in G \\ \emptyset & \text{für } f \in \overline{G}\backslash G \end{cases},$$

genügt es bei Charakterisierungen der Elemente bester Approximation, die Fälle $f \in \overline{G}$ und $f \in R\backslash\overline{G}$ zu unterscheiden (wobei

natürlich nur der zweite Fall interessant ist).

1. Fall: Für eine Teilindexmenge $J \subset I$ gilt: $f \in \overline{G}_\ell$ $(\ell \in J)$. Wegen $f \in \overline{G}_\ell$, $\ell \in J$, und der Voraussetzung $g_\ell \in \mathcal{P}_{G_\ell}(f)$ gilt wegen (1.4) die Aussage

$$(1.5) \quad f = g_\ell \ , \ell \in J.$$

Wegen $\overline{G}_\ell \neq R$ ist die Existenz eines Funktionals $f_\ell^* \in R^*$, $\ell \in J$, mit den Eigenschaften (1.1),(1.2) gesichert und (1.3) ist dann wegen (1.5) automatisch erfüllt.
Die Existenz der übrigen Funktionale f_i^* $(i \in I \backslash J)$ mit den gewünschten Eigenschaften ergibt sich wie im folgenden 2. Fall (dort dann mit $I \backslash J$ statt I).

2. Fall: Es gelte $d := d(f,G_i) > 0$ für $i \in I$.
Aufgrund einer Folgerung aus dem Satz von HAHN-BANACH (siehe etwa unter Folgerung 2 in LJUSTERNIK-SOBOLEV [1965], Seite 109) existiert dann für jedes $i \in I$ ein Element $h_i^* \in R^*$ mit den Eigenschaften

$$(1') \ \|h_i^*\| = 1/d, \ (2') h_i^*(g) = 0 \text{ für } g \in G_i, \ (3') h_i^*(f) = 1.$$

Definiert man nun $f_i^* := \|f - g_i\| \cdot h_i^*$, $i \in I$, so erfüllen diese Funktionale die Behauptung.

-- q.e.d. --

Der vorstehende Satz 1.1 wie auch der nun folgende Satz 1.2 enthält als Spezialfall, nämlich falls $I=\{1\}$ - also bei der Approximation bezüglich nur eines Teilraumes G, Charakterisierungsaussagen für Minimallösungen, die von I.SINGER [1970] - (dort §1.1) - gezeigt wurden. Die dortigen Beweise ließen sich nach entsprechenden Modifikationen auf die uns vorliegende Situation anwenden.

Der folgende Satz 1.2 charakterisiert die Rivlin-Funktion mithilfe des Begriffs des Orthogonalraumes eines Raumes $G \subset R$ in R^*. Zunächst machen wir folgende

Definition 1.2: a) Sei G eine Teilmenge des Raumes R. Dann heißt die Menge

(1.6) $G^{\perp} := \{f^* \in R^* \mid f^*(g) = 0 \text{ für alle } g \in G\}$

Orthogonalraum zu G in R^*. Entsprechend heißt, falls M Teilmenge von R^*,

(1.7) $M_{\perp} := \{f \in R \mid m^*(f) = 0 \text{ für alle } m^* \in M\}$

Orthogonalraum zu M in R.

b) Sei M Teilmenge von R^* und $f \in R$. Dann definieren wir

(1.8) $$\|f\|_M := \sup_{\substack{m^* \in M \\ \|m^*\| \leq 1}} |m^*(f)|$$

Aus der Definition folgt unmittelbar für beliebiges $f \in R$ und $M \subset R^*$ die Gültigkeit von

(1.9) $$\|f\|_M \leq \|f\|$$

Nun können wir folgende Charakterisierungsaussage formulieren:

SATZ 1.2: Die Voraussetzungen seien wie in Satz 1.1. Dann gilt: Es ist $f \in \bigcap_{i \in I} L(g_i)$ genau dann, wenn (1.10) erfüllt ist.

(1.10) Es gilt für alle $i \in I$: $\|f-g_i\|_{G_i^{\perp}} = \|f-g_i\|$

Beweis: Notwendigkeit von (1.10)

Aufgrund von Satz 1.1 existiert für jedes $i \in I$ ein Funktional $f_i^* \in R^*$, welches den Bedingungen (1.1),(1.2) und (1.3) genügt. Dann ist aber

$$\|f-g_i\|_{G_i^{\perp}} \overset{(8)}{=} \sup_{h_i^* \in G_i^{\perp}, \|h_i^*\| \leq 1} |h_i^*(f-g_i)| \geq |f_i^*(f-g_i)| \overset{(3)}{=} \|f-g_i\|$$

für jedes $i \in I$ und zusammen mit $M := G_i^{\perp}$ in (1.9) folgt dann die Behauptung.

Hinlänglichkeit von (1.10)

Es gelte (1.10). Dann ist für jedes $i \in I$

$$\|f-g_i\| \overset{(10)}{=} \|f-g_i\|_{G_i^{\perp}} \overset{(6)}{=} \|f-g\|_{G_i^{\perp}} \overset{(9)}{\leq} \|f-g\|, \quad g \in G_i,$$

somit $g_i \in \mathcal{P}_{G_i}(f)$, also $f \in L(g_i)$.

-- q.e.d. --

Bemerkung 1.1: Unter den Voraussetzungen von Satz 1.1 lassen sich noch eine Reihe äquivalenter Varianten der Bedingungen (1.1),(1.2) und (1.3) angeben. (Vergleiche dazu auch I.SINGER [1970], § 1.1). So ist z.B.: (1.1), (1.2), (1.3) äquivalent zu (1.1), (1.2) und $\mathrm{Re}\, f_i^*(f-g_i) = \|f-g_i\|$, ferner (1.1),(1.2),(1.3) äquivalent zu (1.1),(1.3) und $\mathrm{Re}\, f_i^*(g) = 0$ $(g \in G_i)$.

Ein weiterer Charakterisierungssatz für die Eigenschaft eines Elementes $f \in R$, Rivlin-Funktion bezüglich vorgegebener Elemente $g_i \in G_i \subset R$ $(i \in I)$ zu sein, soll nun im Folgenden angegeben werden. Dabei wird das für normierte lineare Räume verallgemeinerte KOLMOGOROFF-Kriterium (dazu siehe insbesondere BROSOWSKI [1969], BROSOWSKI-WEGMANN [1970]) angewendet auf unser RIVLIN-Problem (P), wobei dann die $G_i \subset R$ keine linearen Teilräume mehr sein müssen, sondern gewisse allgemeinere Teilmengen aus R, nämlich α-Sonnen, sein können.

Definition 1.3: Eine Teilmenge G eines normierten linearen Raumes R heißt eine α-Sonne (in R) genau dann, wenn für jedes Element $f \in R$ jede Minimallösung g_o für f aus G auch eine Minimallösung für sämtliche Elemente der Gestalt $f_\lambda := g_o + \lambda(f-g_o)$ mit $\lambda \geq 1$ ist.

Beispiele für α-Sonnen in R sind lineare Teilräume und konvexe Teilmengen von R (vergl. BROSOWSKI-WEGMANN [1970], Seite 373 ff). Ist $R = C[Q]$ der mit der Tschebyscheff-Norm versehene Vektorraum der auf dem kompakten Hausdorff-Raum Q stetigen reellwertigen Funktionen, so bilden auch die verallgemeinerten rationalen Funktionen (vergl. BROSOWSKI [1968], Seite 90) und, falls $Q = [a,b]$, auch Exponentialsummen der

Form $\sum_{i=1}^{n} a_i e^{b_i x}$ (n fest) Beispiele für α-Sonnen, dagegen Klassen von nichtlinearen Splinefunktionen nicht (vergl. BROSOWSKI [1968], Seite 93).

Definition 1.4: Jedem Element $f \in R$ wird folgende Teilmenge des Dualraums R^* zugeordnet

$$\Sigma_f := \{f^* \in R : \|f^*\| \leq 1 \text{ und } f^*(f) = \|f\|\},$$

die die Menge der Abweichungsfunktionale von f heißt.

Die Menge Σ_f ist konvex und $\sigma(R^*,R)$-kompakt, sie besitzt also Extremalpunkte (siehe etwa KÖTHE [1966], Seite 333 ff). Mit Ep_f bezeichnen wir die Menge dieser Extremalpunkte.

Nach diesen Einführungen können wir folgende Aussage wiedergeben

SATZ 1.3*: (BROSOWSKI-WEGMANN [1970], Seiten 370/1)
Sei G eine α-Sonne in einem normierten Vektorraum R. Sei $f \in R$ und $\tilde{g} \in G$ gegeben. Dann gilt: $\tilde{g}$ ist Minimallösung für f bezüglich G genau dann, wenn (1.11) gilt.

(1.11) Für alle $g \in G$ gilt die Ungleichung

$$\min_{f^* \in Ep_{f-\tilde{g}}} \operatorname{Re} f^*(g-\tilde{g}) \leq 0$$

Bemerkung 1.2: a) Satz 1.3* enthält als Spezialfall (G = linearer Teilraum) den zweiten Charakterisierungssatz von I.SINGER [1970], dort Theorem 1.13 § 1.8 .

b) Ist R = C[Q] (s.o.), so ist ein Funktional f^* aus $\Sigma_{f-\tilde{g}}$ genau dann ein Extremalfunktional, also $f^* \in Ep_{f-\tilde{g}}$, wenn f^* in der Form $\varepsilon(x) f_x^*$ darstellbar ist mit $x \in Q$, $\varepsilon(x) = \operatorname{sgn}\overline{(f(x)-\tilde{g}(x))}$ und $|f(x)-\tilde{g}(x)| = \|f-\tilde{g}\|$, wobei f_x^* das lineare Funktional bezeichnet (→ Punktfunktional, evaluation functional), welches durch $f_x^*(f) = f(x)$ für alle $f \in C[Q]$ erklärt ist. (Dazu siehe etwa SINGER [1970], Seite 69). In diesem Fall geht (1.11) über in das "klassische" KOLMOGOROFF-Kriterium:

(1.11a) Für alle $g \in G$ gilt

$$\min_{\text{über}} (f(x)-\tilde{g}(x)) \cdot (g(x)-\tilde{g}(x)) \leq 0 \ ,$$

wobei das Minimum genommen wird über die Menge der Abweichungspunkte der Fehlerfunktion, d.i. die Menge

$$\{x \in Q \text{ mit } |f(x)-\tilde{g}(x)| = \|f-\tilde{g}\|\}.$$

Aus Satz 1.3* folgt nun unmittelbar ein weiterer Charakterisierungssatz für unser Ausgangsproblem (P):

SATZ 1.3: R sei ein normierter linearer Raum. Die Teilmengen $G_i \subset R$, $i \in I \subset \mathbb{N}$, seien α-Sonnen in R. Seien Elemente $g_i \in G_i$ $(i \in I)$ und $f \in R$ vorgegeben. Dann gilt:

Es ist $f \in \bigcap_{i \in I} L(g_i)$ genau dann, wenn für alle $i \in I$ die Ungleichung

$$\min_{f^* \in Ep_{f-g_i}} \operatorname{Re} f^*(g-g_i) \leq 0 \quad \text{(für alle } g \in G_i)$$

gültig ist.

Ein weiterer Satz charakterisiert eine Minimallösung $\tilde{g} \in G \subset R$ bezüglich $f \in R$ dadurch, daß eine gewisse durch das Differenzelement $f-\tilde{g}$ bestimmte Teilmenge des Dualraumes R^* den Nullvektor aus R^* enthält.

SATZ 1.4*: (BROSOWSKI [1969], Satz 11 und I.SINGER [1969]) Ein Element g_ℓ eines linearen Teilraumes G_ℓ des normierten linearen Raumes R ist genau dann eine Minimallösung für $f \in R$, wenn die $\sigma(G_\ell^*, G_\ell)$-abgeschlossene konvexe Hülle der Menge $\tilde{Ep}_{f-g_\ell}$ den Nullvektor $0^* \in R^*$ enthält, also, abgekürzt geschrieben, wenn gilt

$$0^* \in \operatorname{cl}_{\sigma_\ell} \operatorname{con} \tilde{Ep}_{f-g_\ell}.$$

Dabei bezeichnet $\tilde{Ep}_{f-g_\ell}$ die Menge der auf G_ℓ eingeschränkten linearen Funktionale aus Ep_{f-g_ℓ}.

Hieraus ergibt sich wieder unmittelbar eine Aussage bezüglich des Problems (P):

SATZ 1.4: R sei ein normierter linearer Raum. G_i, $i \in I \subset \mathbb{N}$, seien lineare Teilräume von R. Seien Elemente $g_i \in G_i$, $i \in I$, vorgegeben und $f \in R$. Dann gilt: Es ist $f \in \bigcap_{i \in I} L(g_i)$ genau dann, wenn gilt

$$0^* \in \bigcap_{i \in I} \mathrm{cl}_{\sigma_i} \ \text{con}\ \tilde{E}p_{f-g_i}.$$

(Die Bezeichnungen sind hierbei entsprechend Satz 1.4* gewählt).

Als letzter allgemeiner Charakterisierungssatz für eine Rivlin-Funktion sei in leicht modifizierter Weise ein Satz bewiesen, der von DEUTSCH-MORRIS-SINGER [1969] stammt.

SATZ 1.5: Sei R ein normierter linearer Raum und $\{G_\ell\}$ (ℓ=1,2,...), eine Folge von linearen Teilräumen von R. Ferner seien Elemente $g_\ell \in G_\ell$, (ℓ=1,2,...), gegeben. Dann gilt:
Es existiert ein Element $f \in R$ mit $f \in \bigcap_\ell L(g_\ell)$ genau dann, wenn die zwei folgenden Bedingungen erfüllt sind:

(1.12) Es existieren Elemente $f_\ell \in R$ mit $f_\ell \in A_{G_\ell}^{-1}(0)$, ($\ell$=1,2,...) und

(1.13) es gelten die Beziehungen $g_{\ell+1}-g_\ell = f_\ell - f_{\ell+1}$, ($\ell$=1,2,...)

Beweis: Notwendigkeit der Bedingungen:

Es sei definiert

(1.14) $f_\ell := f-g_\ell$ für ℓ=1,2,..

Aus $f \in \bigcap_\ell L(g_\ell)$ folgt $f \in L(g_\ell)$ für alle ℓ=1,2,...; d.h. $g_\ell \in A_{G_\ell}(f)$, somit

$A_{G_\ell}(f_\ell) = A_{G_\ell}(f-g_\ell) \ni 0$, also $f_\ell \in A_{G_\ell}^{-1}(0)$ für ℓ=1,2,...; d.h. (1.12).

Fortgesetzte Subtraktion von (1.14) ergibt (1.13)

Hinlänglichkeit der Bedingungen:

Mit $f := g_1 + f_1$ folgt aus (1.13), daß gilt $f = f_\ell + g_\ell, \ell \geq 1$. Ferner folgt wegen (1.12)

$\|f-g_\ell\| = \|f_\ell\| = \|f_\ell - 0\| \leq \|f_\ell - (g-g_\ell)\| = \|f-g\|$ für alle $g \in G_\ell$ $(\ell=1,2,..)$
Somit für alle ℓ: $f \in L(g_\ell)$, also $f \in \bigcap_\ell L(g_\ell)$.

-- q.e.d. --

Die vorstehenden Charakterisierungssätze 1.1 - 1.5 machten alle Aussagen von der Art, daß ein Element $f \in R$ dann und nur dann eine RIVLIN-Funktion bezüglich einer Folge $\{g_i\}$, $g_i \in G_i \subset R$ $(i \in I)$ ist, wenn es in Abhängigkeit der gegebenen Elemente g_i und f gewisse Funktionale (Sätze 1.1 - 1.4) beziehungsweise Funktionen (Satz 1.5) gibt, die bestimmten Bedingungen genügen müssen.

Auf der einen Seite sind damit bezüglich unserer Ausgangsfragestellung (P) Charakterisierungen der möglichen Auswahlen von Folgen $\{g_i\}$ bei gewissen gegebenen Teilmengen $G_i \subset R$, $i \in I$, gefunden. Jedoch charakterisieren die vorstehenden Kriterien die "zulässigen" Elemente g_i nur implizit, wobei das Problem der Existenz bzw. Konstruktion von zugehörigen Rivlin-Funktionen gewissermaßen auf das Problem verlagert wird, die Existenz gewisser geeigneter Funktionale oder Funktionen nachzuweisen bzw. sie in Abhängigkeit der vorgegebenen $g_i \in G_i \subset R$ anzugeben (siehe auch Bemerkung 1.3 weiter unten).

Auf der anderen Seite kann man - wie gerade angedeutet - die Charakterisierungssätze dieses Paragraphen als zum Ausgangsproblem (P) äquivalente Problemformulierungen interpretieren, von denen ausgehend man versuchen kann, mehr explizit die eine Rivlin-Funktion $f \in \bigcap_i L(g_i)$ zulassenden Auswahlen von Folgen $\{g_i\}$ zu beschreiben.

Dieser letzten Aufgabe wollen wir uns im weiteren Verlauf dieser Arbeit widmen und dabei insbesondere den Charakterisierungssatz 1.5 im Auge behalten.

Ist nämlich speziell in Satz 1.5 $R = C[a,b]$ der Raum der auf dem endlichen Intervall $[a,b]$ definierten stetigen reellwertigen Funktionen vorgegeben und mit der Tschebyscheff-Norm $\|f\| := \max_{x \in [a,b]} |f(x)|$, $(f \in R)$ - im Folgenden oft auch nur mit T-Norm bezeichnet - versehen, und sind ferner die Teilmengen $G_i \subset R$ Haarunterräume (siehe § 2), so ist der Approximationsoperator A_{G_i} eine eindeutige Abbildung (vergleiche dazu etwa I.SINGER [1970], Seite 103), und die Elemente f_i aus Satz 1.5 lassen sich konstruktiv durch den Alternantensatz (siehe § 2) charakterisieren.

Die Tatsache wird ausgenutzt, um in einigen Fällen eine konkretere Antwort auf die Fragestellung (P) im Raum $C[a,b]$ zu geben.

Dabei spielen im weiteren Verlauf der Untersuchung immer wieder die Begriffe "k-Abweichungsintervall" und "k-Alternierende" eine wichtige Rolle. Diese Begriffe seien deshalb bereits an dieser Stelle eingeführt.

Definition 1.5: a) Eine Funktion $z \in C[a,b]$ heißt k-Nullalternierende, $k \in \mathbb{N}_0$, genau dann, wenn k+2 Punkte x_i existieren, wobei $a \leq x_1 < x_2 < \ldots < x_{k+2} \leq b$, mit der Eigenschaft daß

$$z(x_i) = (-1)^i \varepsilon \|z\|_{[a,b]}, \quad \varepsilon \text{ fest}, \ \varepsilon \in \{-1,+1\} \text{ für } i=1(1)k+2.$$

b) Sei $z \in C[a,b]$ eine k-Nullalternierende und x_i, $i=1(1)k+2$ wie in Teil a) bestimmt. Dann heißt ein Teilintervall

$$S_i^k(z) \subset [a,b], \quad i=1(1)k+2,$$

k-Abweichungsintervall bezüglich z genau dann, wenn gilt

(α) $S_i^k(z)$ enthält x_i und

(β) $S_i^k(z)$ ist Zusammenhangskomponente von x_i in $[a,b]$ mit der Eigenschaft
$z(x) = (-1)^i \varepsilon \|z\|$ für alle $x \in S_i^k(z)$.

(Falls der Bezug eindeutig ist, wird des öfteren auch S_i^k oder $S^k(z)$ statt $S_i^k(z)$ geschrieben und auch vom Abweichungsintervall oder auch von k-Alternierenden gesprochen

c) Ein k-Abweichungsintervall $S_\ell^k(z)$ heißt (-)-Komponente bzw. unteres Abweichungsintervall von z (Bezeichnung: $\underline{S}_\ell^k(z)$), falls $(-1)^\ell \varepsilon = -1$. Entsprechend heißt ein k-Abweichungsintervall $S_\ell^k(z)$ eine (+)-Komponente oder oberes Abweichungsintervall von z (Bezeichnung: $\overline{S}_\ell^k(z)$), falls $(-1)^\ell \varepsilon = +1$.

Bemerkung 1.3: Mit diesen Bezeichnungen sagt unser Satz 1.5 im Fall $R = C[a,b]$, und im Fall, daß G_i ($i \in I$) Haarunterräume sind, aus, daß eine stetige Rivlin-Funktion bezüglich vorgegebener Funktionen $g_i \in G_i$ genau dann existiert, wenn sich i-Nullalternierende f_i ($i \in I$) mit der Eigenschaft (1.13) finden lassen.

Der folgende Paragraph 2 stellt nun mit der Definition von (erweiterten) Tschebyscheff-Systemen und der Beschreibung einiger wichtiger Eigenschaften dieser Funktionenklassen Hilfsmittel für die Paragraphen 3,4,5 zur Verfügung, in denen wir uns mit den weiter oben skizzierten Fragen beschäftigen.

§2 Erweiterte Tschebyscheff-Systeme

- Definition, Eigenschaften und Beispiele -

In den folgenden Paragraphen wird die Problemstellung von RIVLIN untersucht für Funktionen, die sich darstellen lassen als Linearkombinationen aus Funktionen $u_i(x)$, die ihrerseits Tschebyscheff-Systeme bilden. Dabei spielen gewisse Eigenschaften dieser Funktionsklassen eine wichtige Rolle. Im vorliegenden Abschnitt wird nun der Begriff des "erweiterten Tschebyscheff-Systems" definiert, es werden charakteristische Aussagen über diese Systeme und schließlich Beispiele angegeben.

2.1 (E)T-Systeme; Definitionen

Definition 2.1: a) Seien $u_o, u_1, \ldots, u_n$ stetige, reellwertige, auf dem abgeschlossenen endlichen Intervall $[a,b]$ definierte Funktionen. Diese Funktion bilden ein Tschebyscheff-System (kurz: T-System), wenn gilt: Es ist

$$D\begin{pmatrix} 0,1,\ldots,n \\ t_o,t_1,\ldots,t_n \end{pmatrix} := \begin{vmatrix} u_o(t_o) & \ldots & u_o(t_n) \\ \vdots & & \vdots \\ u_n(t_o) & \ldots & u_n(t_n) \end{vmatrix} \neq 0$$

für alle Auswahlen $\{t_\ell\}_o^n$ mit $a \leq t_o < t_1 < \ldots < t_n \leq b$

b) Ein T-System heißt ein vollständiges (complete) T-System (kurz: CT-System), falls die Funktionenmenge $\{u_o, \ldots, u_r\}$ ein T-System bildet für jedes $r \in \{0,1,\ldots,n\}$

Folgerung 2.1: Bilden die Funktionen u_i, $i=0(1)n$, ein T-System, so sind sie linear unabhängig.

finition 2.2: a) Die auf dem abgeschlossenen endlichen Intervall $[a,b]$ definierten reellwertigen Funktionen $u_0,u_1,\ldots,u_n$ bilden ein erweitertes Tschebyscheff-System (kurz: ET-System) der Ordnung z, $z \in \mathbb{N}$, falls gilt: Es ist

1) $\quad u_i \in C^{z-1}[a,b], \quad i=0(1)n \quad$ und

2)

$$D^*\begin{pmatrix} 0,1,\ \ldots,\ n \\ t_0,t_1,\ldots,t_n \end{pmatrix} := \begin{vmatrix} u_0^*(t_0) & \ldots & u_0^*(t_n) \\ \vdots & & \vdots \\ u_n^*(t_0) & \ldots & u_n^*(t_n) \end{vmatrix} \neq 0$$

für alle Auswahlen $\{t_\ell\}_0^n$ mit $a \leq t_0 \leq t_1 \leq \ldots \leq t_n \leq b$, wobei hierbei das Gleichheitszeichen nur gelten darf bei Gruppen von höchstens z aufeinanderfolgenden t_ℓ-Werten. Dabei ist D^* wie folgt definiert:

$u_\ell^*(t_0) := u_\ell(t_0)$ für $\ell=0(1)n$

$u_\ell^*(t_m) := u_\ell(t_m)$ für $\ell=0(1)n$ falls $t_{m-1}<t_m$, $m \in \{1,2,\ldots,n\}$

$u_\ell^*(t_{i+k}) := u_\ell^{(k)}(t_i)$ für $\ell=0(1)n$, $k=0(1)r\leq z-1$ falls

$$t_{i-1}<t_i=t_{i+1}=\ldots=t_{i+r-1}<t_{i+r} \qquad (+)$$

wo $1 \leq i$, $i+r\leq n$ ist und $u_\ell^{(k)}(t_i)$ die k-te Ableitung von $u_\ell(x)$ an der Stelle t_i genommen bedeutet. (Bei i=0 bzw. i+r-1=n entfällt natürlich in (+) der linke resp.rechte Term; vergleiche auch das untenstehende Beispiel zur Bildung von D^*)

b) ET-Systeme der Ordnung n+1 (sonstige Voraussetzungen wie in a)) werden schlechthin als ET-Systeme bezeichnet.

c) Die Funktionen u_ℓ, $\ell=0(1)n$, bilden ein vollständiges ET-System (kurz: ECT-System), falls die Funktionenmenge $\{u_0,u_1,\ldots,u_r\}$ ein ET-System bildet für jedes $r \in \{0,1,\ldots,n\}$.

Beispiel zur Bildung der Determinanten D^* aus Def. 2:
Seien $u_\ell \in C^2[a,b]$, $\ell=0(1)5$, vorgegeben, ferner t_ℓ ($\ell=0(1)5$) mit $a \le t_0 < t_1 = t_2 < t_3 = t_4 = t_5 \le b$.

Dann ist

$$D^*\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ t_0 & t_1 & t_2 & t_3 & t_4 & t_5 \end{pmatrix} = \begin{vmatrix} u_0(t_0) & u_0(t_1) & u_0'(t_1) & u_0(t_3) & u_0'(t_3) & u_0''(t_3) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_5(t_0) & u_5(t_1) & u_5'(t_1) & u_5(t_3) & u_5'(t_3) & u_5''(t_3) \end{vmatrix}$$

Folgerung 2.2: Jedes ET-System (beliebiger Ordnung) ist auch ein T-System.

Dabei ist zu beachten, daß die Umkehrung dieser Aussage nicht gilt.

Definition 2.3: a) Ein (n+1)-dimensionaler Teilraum Q_n von $C[a,b]$ heißt ein (erweiterter) Haar-Unterraum, wenn er eine Basis $\{u_0, \ldots, u_n\}$ besitzt, die ein (E)T-System bildet.

b) Bilden die Funktionen u_ℓ ($\ell=0(1)n$) ein T-System bzw. ET-System, so heißt $Q_n := \operatorname{span}(u_0, \ldots, u_n)$ der zugehörige Haar- (resp. erweiterte Haar-) Unterraum.

c) Ein (E)CT-System heißt auch (erweitertes) Haarsystem.

Definition 2.4: Seien u_ℓ, $\ell=0(1)n$, auf einem Intervall $[a,b]$ definierte Funktionen. Eine Funktion q_n der Form $q_n = \sum_{\ell=0}^{n} c_\ell u_\ell$, $c_\ell \in \mathbb{R}$ heißt u-Polynom oder auch verallgemeinertes Polynom bzgl. der Funktionen u_ℓ, $\ell=0(1)n$. Ein u-Polynom heißt nicht-trivial, falls $\sum_{\ell=0}^{n} c_\ell^2 > 0$ gilt.

Beispielsweise sind die gewöhnlichen algebraischen Polynome $p_n(x) := \sum_{\ell=0}^{n} a_\ell x^\ell$ u-Polynome der speziellen Funktionen $u_\ell(x) := x^\ell$, $0 \le \ell \le n$.

2.2 (E)T-Systeme; Charakterisierung und einige Eigenschaften

Es werden nun bezüglich der im Abschnitt 2.1 definierten Funktionensysteme Charakterisierungen angegeben und zwar mithilfe der durch diese Systeme bildbaren u-Polynome. Ferner wird noch der Alternantensatz und der Eindeutigkeitssatz zitiert, d.h. Aussagen, die die Approximation von Funktionen $f \in C[a,b]$ durch Elemente aus Haar-Unterräumen betreffen. Die unten aufgezählten Sätze erlangen in den folgenden Paragraphen als Hilfsmittel Bedeutung; alle Beweise hierzu kann man z.B. bei KARLIN-STUDDEN [1966] (Seiten 22ff bzw. 284ff) finden.

Zunächst geben wir noch eine kurze Definition.

Definition 2.5: a) Eine isolierte Nullstelle s_0 einer Funktion $f \in C[a,b]$ mit $s_0 \in (a,b)$ heißt G-Nullstelle, falls $f(x)$ in s_0 nicht das Vorzeichen wechselt. Alle anderen isolierten Nullstellen von $f(x)$ in $[a,b]$ einschließlich möglicherweise der Intervallendpunkte heißen U-Nullstellen.

b) Die Anzahl der verschiedenen Nullstellen von f in $[a,b]$ wird mit $Z(f)$ bezeichnet. Die Anzahl der Nullstellen von f in $[a,b]$, wo G-Nullstellen je zweimal und U-Nullstellen einfach gezählt werden, bezeichnet $\tilde{Z}(f)$. Die Anzahl der Nullstellen einer genügend oft differenzierbaren Funktion f in $[a,b]$, wobei die Nullstellen entsprechend ihrer Vielfachheit gezählt werden, wird mit $Z^*(f)$ bezeichnet.

Aufgrund der Definition folgt natürlich sofort, daß gilt: $Z(f) \leq \tilde{Z}(f) \leq Z^*(f)$. Mit den obigen Bezeichnungen gelten folgende Sätze:

SATZ 2.1: a) Sei $(u_\ell,\ \ell=0(1)n)$ ein T-System. Dann gilt $\tilde{Z}(q_n) \leq n$ für jedes nichttriviale u-Polynom q_n. (Hier und in den folgenden Sätzen ist q_n wie in Definition 2.4 bestimmt.)

b) Ist u_ℓ $(\ell=0(1)n)$ eine Menge stetiger Funktionen auf dem Intervall $[a,b]$ und gilt für jedes nichttriviale u-Polynom q_n die Relation $\tilde{Z}(q_n) \leq n$, dann bilden die Funktionen u_ℓ $(\ell=0(1)n)$ ein T-System.

SATZ 2.2: a) Sei $(u_\ell,\ \ell=0(1)n)$ ein ET-System. Dann gilt $Z^*(q_n) \leq n$ für jedes nichttriviale u-Polynom q_n.

b) Gilt $u_\ell \in C^n[a,b]$, $\ell=0(1)n$, und $Z^*(q_n) \leq n$ für alle nichttrivialen u-Polynome q_n, dann bilden die Funktionen u_ℓ, $\ell=0(1)n$, ein ET-System.

SATZ 2.3: Sei $(u_\ell,\ \ell=0(1)n)$ ein T-System und Q_n der zugehörige Haar-Unterraum. Der zugrundeliegende Raum $C[a,b]$ sei mit der T-Norm versehen. Sei $f \in C[a,b]$ vorgegeben.

a) ALTERNANTENSATZ

Für $\tilde{q}_n \in Q_n$ gilt $\|f-\tilde{q}_n\| = \inf_{q \in Q_n} \|f-q\|$ genau dann,

wenn Punkte x_m $(m=1(1)n+2)$ existieren mit $a \leq x_1 < x_2 < \ldots < x_{n+2} \leq b$, sodaß gilt

$f(x_m) - \tilde{q}_n(x_m) = (-1)^i \varepsilon \|f-\tilde{q}_n\|$, $m = 1(1)n+2$,

mit $\varepsilon = +1$ oder $\varepsilon = -1$ als fester Wert.

(Ein Element $\tilde{q}_n$ mit solcher Eigenschaft heißt Minimalpolynom aus Q_n an f oder auch Minimallösung für f bezgl. Q_n)

b) EINDEUTIGKEITSSATZ

Für beliebiges $f \in C[a,b]$ existiert ein Minimalpolynom $\tilde{q}_n \in Q_n$ und es ist eindeutig bestimmt.

2.3 (E)T-Systeme; Beispiele

In den Büchern KARLIN [1968] und KARLIN-STUDDEN [1966] befindet sich eine Fülle von Beispielen für T-Systeme und ET-Systeme. Hier soll nur eine ganz kleine Auswahl wiedergegeben werden, um eine Vorstellung zu gewinnen über mögliche u-Polynome, für die die im Folgenden entwickelten Aussagen der Arbeit gelten.

1. $u_\ell(x) := x^\ell$, $\ell=0(1)n$, $n\in\mathbb{N}$ bilden ein ECT-System. Damit werden die gewöhnlichen algebraischen Polynome erfaßt.

2. Sei $c_0<c_1<\ldots c_n$, $c_\ell\in\mathbb{R}$ für $\ell=0(1)n$. Dann bilden die Funktionen $u_\ell(x) := x^{c\ell}$, $\ell=0(1)n$, ein T-System.

3. Seien die c_ℓ, $\ell=0(1)n$, wie in Beispiel 2. Dann bilden die Funktionen $u_\ell(x) := e^{c\ell x}$, $\ell=0(1)n$, ein ET-System.

4. Seien $g,h\in C^\infty$ und $g'(x)>0$, $h'(x)>0$. Die Funktionen $u_\ell(x) := e^{g(y_\ell)h(x)}$ bilden ein ET-System.

5. Die Funktionen 1, cos t, sin t, cos 2t, sin 2t, ..., cos nt, sin nt mit $0 \le t < 2\pi$ bilden ein T-System (periodisch). In diesem Fall hat das "trigonometrische Polynom der Ordnung n" die Gestalt

$$(2.1)\qquad T_n^*(x) = a_0 + \sum_{\ell=1}^{n} (a_\ell \cos \ell t + b_\ell \sin \ell t);a_\ell,b_\ell\in\mathbb{R}$$

und es gilt hier entsprechend $Z(T_n^*) \le 2n$.

6. Hinreichend dafür, daß ein Funktionensystem $(u_\ell\in C^n,\ell=0(1)n)$ ein ECT-System bildet, ist die Bedingung, daß für alle natürlichen Zahlen r mit $0 \le r \le n$ die Wronski-Determinante $W(u_0(x),\ldots,u_r(x))$ nicht verschwindet auf $[a,b]$, (vergleiche KARLIN [1968], Seite 275).

7. Es sei erwähnt, daß es viele Methoden gibt, aus gegebenen geeigneten Funktionensystemen neue T-Systeme und ET-Systeme zu konstruieren. Ferner läßt sich jedes T-System durch eine glättende Konstruktion in ein ET-System überführen. Näheres dazu findet man z.B. bei KARLIN-STUDDEN [1966], Seite 9ff.

§3 Notwendige Kriterien für die Lösbarkeit bestimmter RIVLIN-Probleme im Raum C[a,b]

Im Folgenden bezeichnet R immer den Raum C[a,b] der stetigen reellwertigen Funktionen auf dem endlichen Intervall [a,b], versehen mit der Tschebyscheff-Norm. Wir betrachten verschiedene Funktionensysteme $F = (u_i,\ i \in I \subset \mathbb{N}) \subset C[a,b]$ und geben uns gewisse aus F bildbare u-Polynome vor. Dabei interessieren wir uns zunächst für notwendige Bedingungen des Rivlin-Problems, also für Kriterien, die erfüllt sein müssen, falls eine Rivlin-Funktion bzgl. der vorgegebenen u-Polynome existiert.

Der erste Satz macht eine Aussage bezüglich der möglichen "Abstände" einer Rivlin-Funktion zu gegebenen u-Polynomen, die aus Haarsystemen gebildet sind.

SATZ 3.1: Die Funktionen $u_\ell \in R$, $\ell = 0(1)n$, mögen ein Haarsystem auf [a,b] bilden. $Q_r := \operatorname{span}(u_0,u_1,\ldots,u_r)$, $r = 0(1)n$, seien die zugehörigen Haar-Unterräume, aus denen u-Polynome der folgenden Gestalt vorgegeben seien:

$$q_r = \sum_{\ell=0}^{r} a_\ell u_\ell \in Q_r \text{ mit } a_r \neq 0.$$

Dann gilt:
Notwendig dafür, daß ein $f \in R$ Element der Menge $L(q_{i_0}, q_{i_1}, \ldots, q_{i_m})$ mit $0 \leq i_0 < \ldots < i_m \leq n$, $0 \leq m \leq n$ ist, ist die Gültigkeit der Ungleichung $E_{i_0}(f) > E_{i_1}(f) > \ldots > E_{i_m}(f)$ für die relevanten Minimalabstände. (Hier und auch später bedeutet $E_r(f) := \inf_{q \in Q_r} \|f-q\|$, ferner ist die Menge $L(q_i)$ und der Approximationsoperator A_{Q_ℓ} wie in §1 definiert.)

Beweis: Es genügt zu zeigen, daß für vorgegebenes $q_s \in Q_s$, $q_t \in Q_t$ mit $0 \leq s < t \leq n$ aus der Existenz einer stetigen Funktion $f \in L(q_s,q_t)$ die Gültigkeit von $E_t(f) < E_s(f)$ folgt. We-

gen $f \in L(q_s,q_t)$ und $Q_s \subset Q_t$ folgt sofort $E_s(f) \geq E_t(f)$. Angenommen, es gelte $E_s(f) = E_t(f)$. $f \in L(q_s,q_t)$ besagt $q_s \in A_{Q_s}(f)$ und $q_t \in A_{Q_t}(f)$. Wegen des Eindeutigkeitssatzes 2.3b und $Q_s \subset Q_t$ würde nun folgen $q_s \equiv q_t$ im Widerspruch zur Auswahl der u-Polynome gemäß der Satzvoraussetzung. Somit gilt die Behauptung.

-- q.e.d. --

Bemerkung 3.1: Für den Spezialfall $u_i(x) := x^i$ hat D.SPRECHER [1969] die obige Aussage bewiesen unter direkter Anwendung des Alternantensatzes. Der vorstehende Beweis mithilfe des Eindeutigkeitssatzes ist kürzer und übersichtlicher.

Im Folgenden wird nun gezeigt, daß das Rivlin-Problem nicht für beliebige u-Polynome aus einem vorgegebenen Funktionensystem lösbar ist, sondern die vorgegebenen u-Polynome z.B. dadurch "gekoppelt" sein müssen, daß sie sich gegenseitig in einer Mindestanzahl von verschiedenen Punkten aus [a,b] schneiden müssen.

SATZ 3.2: Die Funktionen $u_\ell \in R$, $\ell=0(1)n$, mögen ein Haarsystem auf [a,b] bilden, ferner seien u-Polynome $q_r \in Q_r := \text{span}(u_0,u_1,\ldots,u_r)$, $r=0(1)n$, vorgegeben.
Dann gilt folgende Aussage:
Notwendig dafür, daß eine Funktion $f \in R$ Element der Menge $\bigcap_{r=0}^{n} L(q_r)$ ist, ist das Erfülltsein der Bedingung (N):

(N): Für jedes Paar von Indizes (i,ℓ) mit $0 \leq i < \ell \leq n$ gilt entweder $q_\ell - q_i \equiv 0$ oder die u-Polynomdifferenz $q_\ell - q_i$ wechselt das Vorzeichen an mindestens i+1 verschiedenen Punkten des offenen Intervalls (a,b).

Beweis: Es ist $Q_i \subset Q_\ell$. Für $q_i \equiv q_\ell$ ist nichts zu beweisen. Sei also nun $q_i \not\equiv q_\ell$. Wegen der Satzvoraussetzung $f \in \bigcap_{r=0}^{n} L(q_r)$ und dem Vorliegen von Haarunterräumen existieren nach dem Alternantensatz i+2 Punkte aus [a,b] in der Anordnung $a \leq x_1 < x_2 < \ldots$

$\ldots < x_{i+2} \leq b$ mit der Eigenschaft

$$(+) f(x_m) - q_i(x_m) = (-1)^m \varepsilon \, \|f-q_i\|_{[a,b]} = (-1)^m \varepsilon E_i(f),$$

$\varepsilon = +1$ oder -1 fest, $m = 1(1)i+2$

Gemäß Satz 1 gilt $E_i(f) > E_\ell(f)$, folglich

$$\operatorname{sgn}[q_\ell(x_m)-q_i(x_m)] = \operatorname{sgn}\{[f(x_m)-q_i(x_m)] - [f(x_m)-q_\ell(x_m)]\} =$$

$$\operatorname{sgn}[f(x_m)-q_i(x_m)], \quad m = 1(1)i+2$$

Wegen der Stetigkeit der Funktionen und wegen (+) folgt somit die Existenz von mindestens i+1 verschiedenen U-Nullstellen in (a,b).

-- q.e.d. --

Bemerkung 3.2: a) Die vorstehende Behauptung wurde für den Spezialfall $u_\ell(x) = x^\ell$ zuerst von T.J.RIVLIN aufgestellt und bewiesen. Sie ist in der Arbeit von DEUTSCH, MORRIS und SINGER [1969] enthalten.

b) Es sei darauf hingewiesen, daß die Aussage von Satz 2 natürlich auch gültig bleibt für jede Auswahl $(q_{i_0}, q_{i_1}, \ldots, q_{i_m})$ mit $0 \leq i_0 \leq i_1 \leq \ldots \leq i_m \leq n$, $m \leq n$, $i_j \in \{0,1,\ldots,n\}$.

Der folgende Satz schränkt die für die Lösung des Rivlin-Problems geeigneten u-Polynome noch weiter ein, insofern nämlich, als auf bestimmten Teilintervallen von [a,b] ihre Wertevorräte gewissen Mindest- bzw. Höchstschranken unterworfen sein müssen.

SATZ 3.3: Die Funktionen $u_\ell \in R$, $\ell = 0(1)n$ mögen ein Haarsystem auf [a,b] bilden.
Sei $Q_\ell := \operatorname{span}(u_0, u_1, \ldots, u_\ell)$, $\ell = 0(1)n$. Vorgegeben seien u-Polynome $q_i = \sum_{w=0}^{i} a_w u_w \in Q_i$ mit $a_i \neq 0$, $i = r(1)r+t \leq n$, $r \in \mathbb{N}_0$, $t \in \mathbb{N}_0$.

Ferner existiere eine Rivlin-Funktion $f \in \bigcap_{i=r}^{r+t} L(q_i)$.

Unter diesen Voraussetzungen gelten folgende Aussagen:

a) Es existieren Funktionen $f_i \in R$, i=r(1)r+t, die i-Alternierende sind (vergl. Def. 1.5) und für die gilt:

$$(\alpha) \quad f_\ell = f_j - (q_\ell - q_j), \quad r \leq j < \ell \leq r+t$$

$$(\beta) \quad \|f_r\| > \|f_{r+1}\| > \ldots > \|f_{r+t}\|$$

b) Seien $S_\ell^i(f_i)$, ℓ=1(1)i+2, die den i-Alternierenden f_i, i=r(1)r+t, aus Teil a) zugeordneten Abweichungsintervalle (vergl. Def. 1.5). Dann gilt für m=r(1)r+t-1 und i=1(1)r+t-m

$$(\alpha) \quad q_{m+i}(x) - q_m(x) > 0 \text{ für } x \in \bar{S}_j^m(f_m)$$

$$(\beta) \quad q_{m+i}(x) - q_m(x) < 0 \text{ für } x \in \underline{S}_j^m(f_m)$$

c) Mit den Bezeichnungen wie in b) gilt:
Es existiert eine streng monoton fallende Folge von t+1 positiven reellen Konstanten $c_r > c_{r+1} > \ldots > c_{r+t}$, sodaß (mit der Abkürzung $c_{r,s} := c_r - c_s$) folgende Ungleichungen erfüllt sind:

(α) Für alle m=r(1)r+t-1 und i=1(1)r+t-m ist

$$c_{m,m+1} \leq |q_{m+i}(x) - q_m(x)| \text{ falls } x \in S_j^m(f_m), \quad j=1(1)m+2$$

(β) Für alle m=r(1)r+t und s=r(1)m-1 ist

$$q_m(x) - q_s(x) \leq c_{s,m} \quad \text{falls} \quad x \in \bar{S}_j^m(f_m)$$

(γ) Für alle m=r(1)r+t und s=r(1)m-1 ist

$$-(q_m(x) - q_s(x)) \leq c_{s,m} \text{ falls } x \in \underline{S}_j^m(f_m)$$

Beweis: zu a): Die Existenz von i-Alternierenden mit der Eigenschaft (α) folgt unmittelbar aus Satz 1.5 und Satz 2.3

(siehe auch Bemerkung 1.3). Ebenso entnimmt man dem Beweis von Satz 1.5 die Gleichung $f = f_s + q_s$ für $s = r(1)r+t$. Wegen $0 \in A_{Q_s}^{-1}(f_s)$ folgt dann $E_s(f) = \|f_s\|$ und nach Satz 3.1 schließlich (β).

zu b): (α): Sei x irgendein Punkt eines oberen Abweichungsintervalles $\bar{S}_j^m(f_m)$, d.h. es gilt $f_m(x) = \|f_m\|$. Dann ist wegen Teil a (α) und der Definition der T-Norm

$$\|f_{m+i}\| + q_{m+i}(x) - q_m(x) \geq f_{m+i}(x) + q_{m+i}(x) - q_m(x) = f_m(x) = \|f_m\| .$$

Wegen Teil a (β) ist somit $q_{m+i}(x) - q_m(x) \geq \|f_m\| - \|f_{m+i}\| > 0$. Also gilt (α). Behauptung (β) folgt ähnlich: Angenommen, es gelte für einen Punkt $\tilde{x}$ eines unteren Abweichungsintervalles $\underline{S}_j^m(f_m)$ die Ungleichung $q_{m+i}(\tilde{x}) - q_m(\tilde{x}) \geq 0$. Dann ist

$$-\|f_{m+i}\| \leq f_{m+i}(\tilde{x}) = f_m(\tilde{x}) - (q_{m+i}(\tilde{x}) - q_m(\tilde{x})) \leq f_m(\tilde{x}) = -\|f_m\| .$$

Somit wäre $\|f_m\| \leq \|f_{m+i}\|$. Dies ist aber ein Widerspruch zu Behauptung a). Somit gilt $q_{m+i}(x) - q_m(x) < 0$ für alle $x \in \underline{S}_j^m(f_m)$, also die Behauptung.

zu c): (α): Die Abweichungsintervalle sind entweder (-)-Komponenten oder (+)-Komponenten. Die Behauptung wird für die Komponenten einzeln gezeigt. Sei also zunächst x beliebiger Punkt aus $\bar{S}_j^m(f_m)$. Wegen b (α) ist $q_{m+i}(x) - q_m(x) > 0$, somit $|q_{m+i}(x) - q_m(x)| = q_{m+i}(x) - q_m(x)$. Es ist

$$(+)\quad \|f_{m+i}\| \geq f_{m+i}(x) = f_m(x) - (q_{m+i}(x) - q_m(x)) = \|f_m\| - |q_{m+i}(x) - q_m(x)| .$$

Sei nun x beliebig aus $\underline{S}_j^m(f_m)$. Wegen b (β) ist $q_{m+i}(x) - q_m(x) < 0$, ferner

$$(++)\quad -\|f_{m+i}\| \leq f_{m+i}(x) = -\|f_m\| - (q_{m+i}(x) - q_m(x)) = -\|f_m\| + |q_{m+i}(x) - q_m(x)|$$

Die Ungleichungen (+) und (++) ergeben die Gültigkeit von $\|f_m\| - \|f_{m+i}\| \leq |q_{m+i}(x) - q_m(x)|$ für jeden Punkt eines Abweichungsintervalles $S_j^m(f_m)$. Setze $c_k := \|f_k\|$, $k = r(1)r+t$.

Zu (β): Sei $x \in \bar{S}_j^m(f)$, dann $\|f_s\|-\|f_m\| \geq f_s(x)-\|f_m\| = f_s(x)-f_m(x) = q_m(x)-q_s(x)$.

Zu (γ): Sei $x \in \underline{S}_j^m(f)$, dann $-\|f_m\| = f_m(x) = f_s(x)-(q_m(x)-q_s(x)) \geq -\|f_s\|-(q_m(x)-q_s(x))$.

-- q.e.d. --

Bemerkung 3.3: Wie aus dem vorstehenden Beweis hervorgeht, gilt die Satzaussage nicht nur bei "direkt aufeinanderfolgenden" u-Polynomen q_i, $i=r(1)r+t \leq n$. Die Sequenz der vorgegebenen u-Polynome muß nur bezgl. der Indizes streng monoton aufsteigend geordnet sein, also z.B. können gegeben sein q_i mit $i=r(t_\ell)r+t \leq n$, wo $t_\ell \in \mathbb{N}$ mit $t_\ell \leq t$.

Für den Fall, daß ein ECT-System vorliegt und die Rivlin-Aufgabe für ein Tripel "direkt aufeinanderfolgender" (vergl. unten) u-Polynome betrachtet wird, kann man die in Satz 3.2 getroffene Aussage über die notwendige Mindestanzahl von Schnittpunkten gewisser Differenzenpolynome dahingehend verfeinern, daß bei vorausgesetzter Existenz einer Rivlin-Funktion sogar bestimmte eindeutige Zuordnungen dieser Schnittstellen bestehen müssen.

SATZ 3.4: Die Funktionen $u_\ell \in R$, $\ell=0(1)n \geq 2$, mögen ein erweitertes Haarsystem auf dem Intervall $[a,b]$ bilden. Sei $Q_\ell := \mathrm{span}(u_0,u_1,\ldots,u_\ell)$, $\ell=0(1)n$. Gegeben sei ein u-Polynom-Tripel$_i(q_r,q_{r+1},q_{r+2})$ mit $0 \leq r$, $r+2 \leq n$ von u-Polynomen $q_i = \sum_{s=0}^{i} a_s u_s \in Q_i$ mit $a_i \neq 0$, $i=r(1)r+2$.
Es existiere eine Rivlin-Funktion $f \in L\,(q_r,q_{r+1},q_{r+2})$.
Dann gelten folgende Aussagen:

a) Die Differenzfunktionen $q_{r+1}-q_r$, $q_{r+2}-q_r$ und $q_{r+2}-q_{r+1}$ besitzen jeweils nur lauter einfache U-Nullstellen im Innern von $[a,b]$.

b) Seien mit x_i die Nullstellen in [a,b] von $q_{r+2}(x)-q_r(x)$ bezeichnet und mit y_m die Nullstellen in [a,b] von $q_{r+1}(x)-q_r(x)$, jeweils der Größe nach aufsteigend angeordnet.
Dann gibt es eine eindeutige paarweise Zuordnung der Nullstellen x_i und y_m in der Form $x_i \leftrightarrow y_i$ (i=1(1)r+1) oder in der Form $x_{i+1} \leftrightarrow y_i$(i=1(1)r+1) mit der Eigenschaft, daß die zu den Nullstellen gehörigen Differenzpolynome in den sich entsprechenden Nullstellenpaaren gleichsinnig ihr Vorzeichen wechseln.

c) Je nach Zuordnung gemäß Teil b) gilt

falls $Z(q_{r+2}-q_r)=r+1$: $\max(x_i,y_i)<\min(x_{i+1},y_{i+1})$ für i=1(1)r

falls $Z(q_{r+2}-q_r)=r+2$: es ist

entweder $\max(x_i,y_i)<\min(x_{i+1},y_{i+1})$ für i=1(1)r und $x_{r+2}>\max(x_{r+1},y_r)$

oder $\max(x_{i+1},y_i)<\min(x_{i+2},y_{i+1})$ für i=1(1)r und $x_1<\min(x_2,x_1)$

Beweis: zu a): Gemäß Satz 3.2 hat $q_i(x)-q_\ell(x)$, $r\leq\ell<i\leq r+2$ mindestens $\ell+1$ verschiedene U-Nullstellen in (a,b). Da ein ECT-System zugrunde liegt, gilt gemäß Satz 2.2 $Z^*(q_i-q_\ell)\leq i$ (*), $r\leq\ell<i\leq r+2$.
Bleibt zu betrachten die Funktion $q_{r+2}(x)-q_r(x)$: Wäre unter den nach Satz 3.2 existierenden r+1 verschiedenen U-Nullstellen dieser Funktion eine mehrfache Nullstelle, so müßte sie als U-Nullstelle eine Vielfachheit $v \geq 3$ haben mit $v \equiv 1 \mod (2)$. Das ergäbe aber einen Widerspruch zu (*). Ebenfalls wegen (*) könnte, falls $Z(q_{r+2}-q_r) = r+2$, die Vielfachheit jeder Nullstelle genau nur 1 sein.

zu b): Gemäß Satz 3.3 existieren ℓ-Alternierende f_ℓ (ℓ=r,r+1,r+2) zu den drei vorgegebenen u-Polynomen q_ℓ.

Insbesondere also existieren r+2 Punkte s_i :
$a \leq s_1 < s_2 < \ldots < s_{r+2} \leq b$ mit

$$(\dagger) \qquad f_r(s_i) = (-1)^i \varepsilon \|f_r\|, \ \varepsilon = \pm 1 \text{ fest}, \ i = 1(1)r+2.$$

Sei o.B.d.A. $q_{r+1}(t) - q_r(t) > 0$ für $a \leq t < y_1$. (Alle Nullstellen von $q_{r+1}(t) - q_r(t)$ liegen ja gemäß Satz 2.2 und Satz 3.2 im Innern von $[a,b]$). Wegen der Stetigkeit von $q_{r+1} - q_r$ und Satz 3.3(b) gilt dann

$$(\dagger\dagger) \qquad a \leq s_1 < y_1 < s_2 < y_2 < \ldots < y_{r+1} < s_{r+2} \leq b.$$

und in (†) gilt $\varepsilon = +1$. Die Funktion $q_{r+2} - q_r$ besitzt gemäß Satz 3.2 und Teil a) dieses Satzes mindestens r+1 einfache Nullstellen in (a,b), insgesamt höchstens r+2 Nullstellen. Wegen Satz 3.3(b), wegen (†) und der Stetigkeit von $q_{r+2} - q_r$ muß in jedem offenen Intervall (s_{i-1}, s_i), i=2(1)r+2, mindestens eine Nullstelle von $q_{r+2}(t) - q_r(t)$ liegen. Es liegt sogar jeweils genau eine dort:

1. Fall: $Z(q_{r+2} - q_r) = r+1$. Dann ist $\{x_i, y_i\} \subset (s_i, s_{i+1})$, i=1(1)r+1, und es gilt die Behauptung.
2. Fall: Es sei nun $Z(q_{r+2} - q_r) = r+2$. Lägen in einem Intervall $(s_{\ell_o - 1}, s_{\ell_o})$, $\ell_o \in \{2,3,\ldots,r+2\}$, zwei Nullstellen von $q_{r+2} - q_r$, so wären sie wegen ihrer Einfachheit U-Nullstellen und es würde gelten

$$(\dagger\dagger\dagger) \qquad \operatorname{sgn}(q_{r+2}(s_{\ell_o}) - q_r(s_{\ell_o})) = -\operatorname{sgn}(q_{r+1}(s_{\ell_o}) - q_r(s_{\ell_o})),$$

da nicht drei einfache Nullstellen von $q_{r+2} - q_r$ zwischen benachbarten Abweichungspunkten von f_r liegen können. (†††) ist aber ein Widerspruch zu Satz 3.3(b). Insgesamt gilt also dann entweder $x_1 < s_1$ und $\{x_{i+1}, y_i\} \subset (s_i, s_{i+1})$ oder $\{x_i, y_i\} \subset (s_i, s_{i+1})$ und $s_{r+2} < x_{r+2}$ für i=1(1)r+1. Die Paarbildung ist somit eindeutig und durch die Lage von x_1 bestimmt.
Die Gleichsinnigkeit des Nulldurchlaufs folgt wiederum aus (†) und Satz 3.3(b).

zu c): Die Behauptung ergibt sich aus dem vorstehenden Beweisgang zu Teil a) und b) unmittelbar.

Bemerkungen zu SATZ 3.4 und Folgerungen:

. Bedingung c) von Satz 3.4 schließt für ein u-Polynomtripel, welches eine Rivlin-Funktion zuläßt, Kurvenverläufe von der Art aus, die Bild 1 qualitativ andeutet:

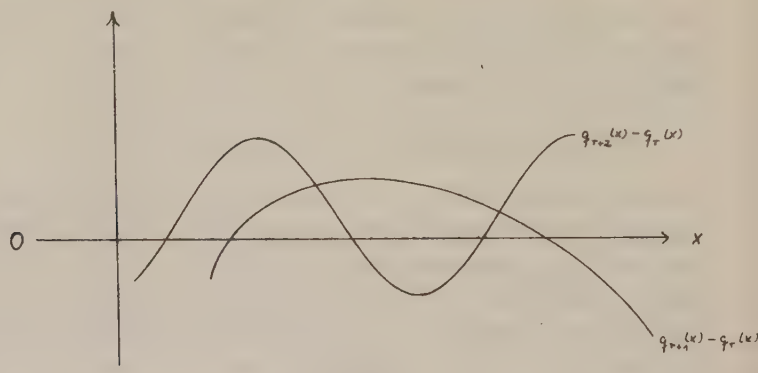

BILD 1

2. a) Satz 3.4 macht nur eine Aussage über die paarweise Zuordnung, nicht aber darüber, welches Element eines Paares das kleinere bzw. größere ist. Dabei wird auch nicht ausgeschlossen, daß die zwei Elemente eines solchen Nullstellenpaares zusammenfallen.
   b) Sei {x,y} ein zusammengehöriges Nullstellenpaar im Sinne von Satz 3.4(b). Gilt x=y, so ist wegen $q_{r+2}-q_{r+1}=(q_{r+2}-q_r)-(q_{r+1}-q_r)$ dieser Punkt auch zugleich Nullstelle von $q_{r+2}(t)-q_{r+1}(t)$. Gilt $x \neq y$, so kann dann im Intervall $[\min(x,y), \max(x,y)]$ keine Nullstelle z der Funktion $q_{r+2}(t)-q_{r+1}(t)$ liegen.

3. Skizziert man die möglichen Anordnungen der Nulldurchgänge der Differenzenfunktionen $q_{r+2}-q_r$ und $q_{r+1}-q_r$ zusammen mit den zugehörigen Tangentenrichtungen qualitativ, so ergeben sich aufgrund Satz 3.4(b) folgende vier Grundtypen, die in Bild 2 am Beispiel r=3 angedeutet werden:

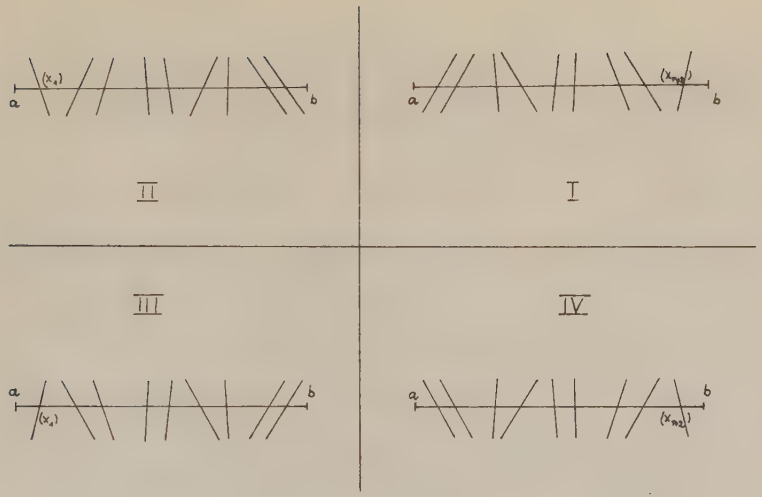

BILD 2

Hierbei kann x_1 bei den Typen II, III resp. x_{r+2} bei den Typen I, IV eventuell wegfallen, nämlich in dem Fall $Z(q_{r+2}-q_r) = r+1$.
Da jeder der Typen I - IV aus jedem anderen durch geeignete Spiegelungen und Transformationen hervorgeht, kann man sich bei vielen weiteren Betrachtungen o.B.d.A. auf nur einen der Grundtypen beschränken.

Zum Abschluß dieses Paragraphens wird noch gezeigt, daß die Auswahl von u-Polynomtripeln, die eine Rivlin-Funktion zulassen, noch dadurch weiter eingeschränkt wird, daß an genau einem Intervallendpunkt R_1 endweder die Relation $q_{r+2}(R_1) < q_{r+1}(R_1) < q_r(R_1)$ oder die Relation $q_{r+2}(R_1) > q_{r+1}(R_1) > q_r(R_1)$ gelten muß. Diese Aussage ist enthalten in folgendem

SATZ 35: Es mögen die Voraussetzungen von Satz 34 gelten. Dann gilt:
An genau einem Randpunkt $R_1 \in \{a,b\}$ haben alle drei Funktionen $q_{r+2}(x)-q_r(x)$, $q_{r+2}(x)-q_{r+1}(x)$ und $q_{r+1}(x)-q_r(x)$ gleiches Vorzeichen.

Beweis: Unter den Voraussetzungen gilt $Z(q_{r+2}-q_{r+1}) = Z^*(q_{r+2}-q_{r+1}) = r+2$, ebenso $Z^*(q_{r+1}-q_r) = Z(q_{r+1}-q_r) = r+1$ und alle Nullstellen sind einfache U-Nullstellen im Innern von [a,b]. Da ferner $Z(q_{r+2}-q_{r+1}) = Z(q_{r+1}-q_r) + 1$ gilt, folgt notwendig:

(†) $\operatorname{sgn}\{q_{r+2}(R_1)-q_{r+1}(R_1)\} = \operatorname{sgn}\{q_{r+1}(R_1)-q_r(R_1)\}$ für entweder $R_1=a$ oder $R_1=b$ und

$\operatorname{sgn}\{q_{r+2}(R_2)-q_{r+1}(R_2)\} = -\operatorname{sgn}\{q_{r+1}(R_2)-q_r(R_2)\}$ für $R_2 \in \{a,b\}$ mit $R_2 \neq R_1$.

Weiterhin gilt $q_{r+2}(x)-q_{r+1}(x)>0$ (bzw.<0) genau dann, wenn auch gilt $q_{r+2}(x)-q_r(x) > q_{r+1}(x)-q_r(x)$ (bzw. $< q_{r+1}(x)-q_r(x)$). Zusammen mit (†) folgt $\operatorname{sgn}\{q_{r+2}(R_1)-q_r(R_1)\} = \operatorname{sgn}\{q_{r+2}(R_1)-q_{r+1}(R_1)\} = \operatorname{sgn}\{q_{r+1}(R_1)-q_r(R_1)\} = \varepsilon$ mit $\varepsilon \in \{1,-1\}$.

-- q.e.d. --

Bemerkungen zu SATZ 3.5 und Folgerungen:

1. Voraussetzungen wie in Satz 3.5. Dann können folgende zwei Bedingungen nicht gleichzeitig für t=a und t=b gelten:
   a) $\operatorname{sgn}\{q_{r+2}(t)-q_r(t)\} = \operatorname{sgn}\{q_{r+1}(t)-q_r(t)\}$
   b) $|q_{r+2}(t)-q_r(t)| > |q_{r+1}(t)-q_r(t)|$

   (Diese Aussage bleibt gültig, wenn ">" durch "<" ersetzt wird.)

2. Der in Satz 3.5 ausgezeichnete Randpunkt R_1 ist bereits durch die Werte der vorgegebenen u-Polynome an den Intervallendpunkten eindeutig bestimmt.

3. Es gilt $Z(q_{r+2}-q_r)=r+2$ genau dann, wenn die Gleichung $\operatorname{sgn}\{q_{r+2}(t)-q_r(t)\} = \operatorname{sgn}\{q_{r+1}(t)-q_r(t)\}$ nicht zugleich für t=a und t=b besteht.

4. Gilt $Z(q_{r+2}-q_r)=r+2$, so gilt für die Grundtypen I und IV, daß $R_1=a$, ferner für die Grundtypen III und II, daß $R_1=b$ ist (vergleiche Bemerkungen zu Satz 3.4 und Bild 2).

§4 Das RIVLIN-Problem in C[a,b]: Einige Charakterisierungssätze

Während im vorangegangenen Paragraphen eine Reihe von notwendigen Bedingungen für die Existenz einer Rivlin-Funktion bzgl. gewisser gegebener u-Polynome angegeben wurde, sollen nun spezielle Charakterisierungssätze zum Rivlin-Problem im Raum R := C[a,b] - versehen mit der Tschebyscheff-Norm - gezeigt werden. Dabei werden zum Beweis der Hinlänglichkeit gewisser Bedingungen für die Existenz einer Rivlin-Lösung durchweg konstruktive Methoden angewendet.

Der erste Satz verallgemeinert Ergebnisse von DEUTSCH, MORRIS, SINGER [1969] und D.SPRECHER [1969]. Erstere zeigten für den Speziallfall $u_o(x)=1$ und $u_1(x)=x$, daß die notwendige Bedingung aus Satz 3.2 auch hinreichend ist für die Existenz einer stetigen Rivlin-Funktion $f \in L(p_o,p_1)$, $p_o \in \text{span}(1)$, $p_1 \in \text{span}(1,x)$. D.SPRECHER zeigte die Gültigkeit dieser Tatsache für beliebige Paare (p_n,p_m) von algebraischen Polynomen.
Der unten stehende neue Beweis ist kürzer und durchsichtiger als der in der Arbeit von D.SPRECHER [1969]. Das hier entwickelte Konstruktionsprinzip, welches mit dem Satz 3.3(a) bzw. dem Satz 1.5 eng verknüpft ist, läßt sich auch auf komplexere Fälle ausdehnen, die z.B. in Satz 4.2 und Satz 4.3 auftreten.

SATZ 4.1: Die Funktionen $u_\ell \in R$, $\ell=0(1)n$, mögen ein Haarsystem auf dem Intervall [a,b] bilden. Mit Q_r, $r=0(1)n$, seien die von u_ℓ, $\ell=0(1)r$, erzeugten Haarunterräume bezeichnet. Seien verallgemeinerte Polynome $q_m \in Q_m$ und $q_t \in Q_t$, $0 \leq m < t \leq n$, mit $q_t \neq q_m$ vorgegeben. Dann gilt: Es existiert genau dann ein $f \in R$ mit $f \in L(q_t,q_m)$, wenn die Funktion $q_t - q_m$ mindestens m+1 verschiedene U-Nullstellen im Innern von [a,b] besitzt.

Beweis: A: Die Notwendigkeit der Bedingung ergibt sich aus Satz 3.2

B: Beweis der Hinlänglichkeit der Bedingung

1. vorbereitende Bezeichnungen

Mit x_i $(i=1,\ldots,k \geq m+1)$ seien die k Stellen bezeichnet, in denen die Funktion $d(x) := q_t(x) - q_m(x)$ in $[a,b]$ das Vorzeichen wechselt. Die Anordnung dieser U-Nullstellen sei $x_0 := a < x_1 < x_2 < \ldots < x_k < b$. Das Intervall $[a,b]$ werde wie folgt unterteilt: Es sei $\tilde{B}_{\ell-1} := [x_{\ell-1}, x_\ell]$ für $\ell=1(1)m+1$ und $\tilde{B}_{m+1} := [x_{m+1}, b]$. Nun wähle Teilintervalle $B_{\ell-1} := [\alpha_{\ell-1}, \beta_{\ell-1}] \subset \tilde{B}_{\ell-1}$, $\ell=1(1)m+2$, mit folgenden drei Eigenschaften:

(a) $x_{\ell-1} < \alpha_{\ell-1} < \beta_{\ell-1} < x_\ell$ für $\ell=1(1)m+1$

(b) $x_{m+1} < \alpha_{m+1} < \beta_{m+1} < b$ falls $k=m+1$

$x_{m+1} < \alpha_{m+1} < \beta_{m+1} < x_{m+2}$ falls $k>m+1$

(c) $d(x) \neq 0$ für alle $x \in B_\ell$, $\ell=0(1)m+1$

Wegen $d \not\equiv 0$ existieren solche Teilintervalle immer. Nun werde noch eine konstante c_m gewählt, die der Bedingung $c_m > \|d\|$ genügt.

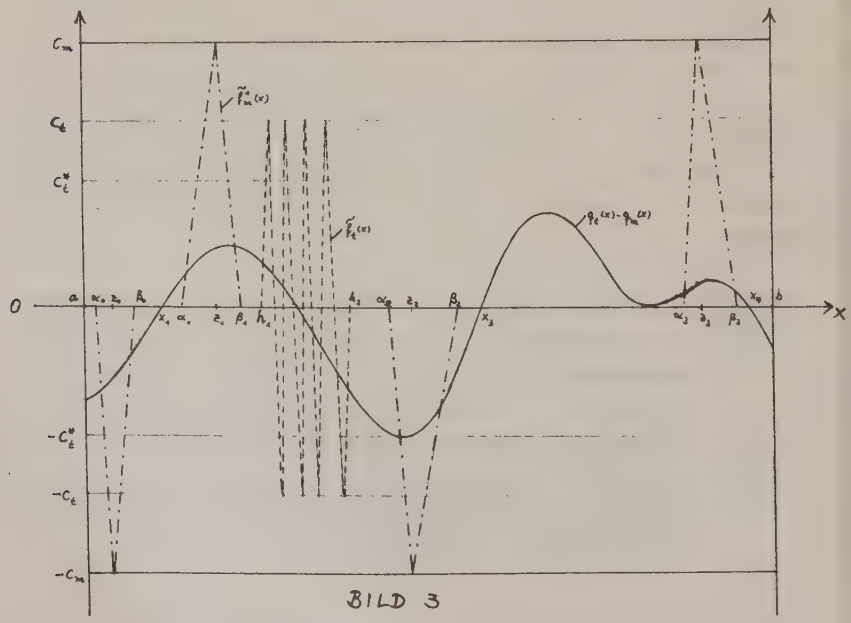

BILD 3

Beispiel zur Beweiskonstruktion mit $t=6$, $m=2$ und $k=4$

2. Konstruktion einer m-Nullalternierenden

Sei z_ℓ ein beliebiger, fest gewählter innerer Punkt aus B_ℓ, $\ell=0(1)m+1$. Es wird nun eine Funktion $\tilde{f}_m^\ell(x)$ auf $\tilde{B}_\ell$, $\ell=0(1)m+1$, wie folgt definiert:

$$\tilde{f}_m^\ell(x) := \begin{cases} c_m \cdot \operatorname{sgn} d(x) \cdot (x-\alpha_\ell)/(z_\ell-\alpha_\ell) & \text{für } \alpha_\ell \le x \le z_\ell \\ c_m \cdot \operatorname{sgn} d(x) \cdot (x-\beta_\ell)/(z_\ell-\beta_\ell) & \text{für } z_\ell \le x \le \beta_\ell \\ 0 & \text{sonst in } \tilde{B}_\ell \end{cases}$$

Da $\tilde{B}_\ell \cap \tilde{B}_{\ell+1} = \{x_{\ell+1}\}$, $\ell=0(1)m$, und $\tilde{f}_m^\ell(x_{\ell+1}) = \tilde{f}_m^{\ell+1}(x_{\ell+1})$ gilt, ist durch

$$\tilde{f}_m(x) := \tilde{f}_m^\ell(x) \quad \text{falls } x \in \tilde{B}_\ell$$

auf dem ganzen Intervall $[a,b]$ eine stetige Funktion definiert, denn $\bigcup_{\ell=0}^{m+1} \tilde{B}_\ell$ überdeckt $[a,b]$. Weiter wird für $x \in [a,b]$ gesetzt:

$$f_m^*(x) := \begin{cases} \max \{d(x), \tilde{f}_m(x)\} & \text{für } x \text{ mit } \operatorname{sgn} d(x)=+1 \\ \min \{d(x), \tilde{f}_m(x)\} & \text{für } x \text{ mit } \operatorname{sgn} d(x)=-1 \\ 0 & \text{sonst} \end{cases}$$

Nach Konstruktion ist $f_m^* \in R$, ferner gilt $f_m^*(z_\ell)=c_m \cdot \operatorname{sgn} d(z_\ell)= (-1)^\ell \varepsilon c_m$ mit $\varepsilon = +1$ oder $\varepsilon = -1$, und $\ell=0(1)m+1$. Ansonsten gilt $|f_m^*(x)| \le c_m$ für alle $x \in [a,b]$. Somit ist $\|f_m^*\| = c_m$ und nach dem Alternantensatz 2.3(a) ist das identisch verschwindende u-Polynom $0 \in Q_m$ Minimallösung für f_m^* bzgl. Q_m.

3. Konstruktion einer t-Nullalternierenden

Sei nun definiert $g := \min_{0 \le \ell \le m+1} \{ \min_{x \in B_\ell} (c_m - |d(x)|)\}$.

Wegen $B_\ell \subset \tilde{B}_\ell$ ist lt. Konstruktion $g > 0$ und $c_t^* := c_m - g > 0$.

Mit diesen Bezeichnungen gilt nach Teil 1: $\|f_m^* - d\| \le c_t^*$.

Wähle c_t mit $c_t^* \le c_t < c_m$ und fixiere irgendeinen Nulldurchgang x_s von $d(x)$, $s \in \{1,\dots,m+1\}$.

Wegen der Stetigkeit von d(x) läßt sich eine Umgebung $U := U(x_s) = [h_1,h_2]$ angeben mit den zwei Eigenschaften

$$\text{(a)} \quad \beta_{s-1} \leq h_1 < x_s < h_2 \leq \alpha_s$$

$$\text{(b)} \quad \|q_t - q_m\|_U = \|d\| \leq c_m - c_t$$

In U wähle man Punkte r_i, i=1(1)t+2, mit $h_1 < r_1 < \ldots < r_{t+2} < h_2$. Auf [a,b] wird nun folgende Funktion definiert

$$\tilde{f}_t(x) := \begin{cases} c_t(x-h_1)/(r_1-h_1) & \text{für } x \in [h_1,r_1] \\ c_t(r_{2\ell-1}+r_{2\ell}-2x)/(r_{2\ell}-r_{2\ell-1}) & \text{für } x \in [r_{2\ell-1},r_{2\ell}], \\ & \ell=1(1)[(t+2)/2] \\ c_t(2x-r_{2\ell}-r_{2\ell+1})/(r_{2\ell+1}-r_{2\ell}) & \text{für } x \in [r_{2\ell},r_{2\ell+1}], \\ & \ell=1(1)[(t+1)/2] \\ (-1)^{t+1} c_t(h_2-x)/(h_2-r_{t+2}) & \text{für } x \in [r_{t+2},h_2] \\ 0 & \text{sonst} \end{cases}$$

Man sieht unmittelbar, daß gilt $\|\tilde{f}_t\| = c_t$ und daß $\tilde{f}_t \in A_{Q_t}^{-1}(0)$.

4. Definition der gesuchten Rivlin-Funktion

$$\text{Sei } f_m(x) := \begin{cases} f_m^*(x) & \text{für } x \in [a,b]-U \\ \tilde{f}_t(x) + (q_t(x)-q_m(x)) & \text{für } x \in U \end{cases}$$

Wegen $\|q_t - q_m\|_U \leq c_m - c_t$ gilt auch $\|f_m\|_{[a,b]} = c_m$ und $f_m \in A_{Q_m}^{-1}(0)$.

Sei $f_t(x) := f_m(x) - (q_t(x)-q_m(x))$ für $x \in [a,b]$

Es ist $|f_t(x)| \leq c_t^* \leq c_t$ für $x \in [a,b]-U$ gemäß Teil 2 und $f_t(x) = \tilde{f}_t(x)$ für $x \in U$.

Somit gilt $\|f_t\|_{[a,b]} = c_t$ und $f_t \in A_{Q_t}^{-1}(0)$.

Schließlich werde definiert

$$f(x) := f_m(x) + q_m(x) = f_t(x) + q_t(x)$$

Nach Konstruktion ist $f \in R$. Daß gilt: $f \in L(q_t) \cap L(q_m)$, ergibt sich nun direkt aus den Eigenschaften von f_m und f_t sowie aus Satz 2.3(a).

-- q.e.d. --

Bemerkung 4.1: Aus dem Beweisgang ist unmittelbar ersichtig, daß eine Rivlin-Funktion nicht eindeutig bestimmt ist.

Satz 4.1 besagte, daß im Falle eines vorgegebenen u-Polynompaares die notwendige Bedingung des Satzes 3.2 auch hinreichend ist für die Existenz einer stetigen Rivlin-Funktion f, und in dieser Form war die Aussage eine Übertragung des Ergebnisses von D.SPRECHER [1969] auf das Vorliegen von Haarsystemen. Die Bedingung von Satz 4.1 reicht aber sogar dafür aus, die Klasse der möglichen Rivlin-Funktionen in Abhängigkeit von Eigenschaften der vorgegebenen u-Polynome noch besser zu charakterisieren. Genauer gesagt, gilt:

SATZ 4.2: Es mögen die Voraussetzungen von Satz 4.1 gelten und es gelte zusätzlich
$q_t \in C^h[a,b]$ und $q_m \in C^\ell$, $h \in \mathbb{N}_0 \geq \ell$, ($h,\ell = \infty$ zugelassen)
Dann gilt:
Es existiert eine Funktion $f \in C^{\min(h,\ell)}[a,b]$ mit $f \in L(q_m,q_t)$ genau dann, wenn $q_t - q_m$ mindestens m+1 verschiedene U-Nullstellen im Innern von [a,b] hat.

Beweis: A. Notwendigkeit: Wegen $C^{\min(h,\ell)} \subset C^0$, folgt die Beh. aus Satz 3.2
B. Beweis der Hinlänglichkeit der Bedingung:

1. Vorbereitendes

Die Beweisstruktur ist ähnlich dem Beweis von Satz 4.1. Für die Definition der Nullalternierenden wird jedoch noch folgendes Lemma benötigt: (vergl. DE RHAM [1955], Seite 4)

Lemma: Sei M_1 eine kompakte Teilmenge aus $\mathbb{R}^n$ und M_3 eine M_1

umfassende offene Teilmenge. Dann existiert eine auf M_3 definierte Funktion F mit kompaktem Träger M_2, wobei $M_1 \subset M_2 \subset M_3$, mit folgenden Eigenschaften

(a) $F \in C^\infty[M_3]$

(b) $F(x) = 1$ für $x \in M_1$

(c) $0 \leq F(x) \leq 1$ für $x \in M_2 - M_1$

Die Größen x_i (i=0(1)k), $\tilde{B}_i$ und B_i (i=0(1)m+1) sowie c_m seien genau wie im Beweis zu Satz 4.1 Teil B1. definiert.

2. Konstruktion einer m-Alternierenden

Sei $Z_i := [v_i, b_i] \subset B_i$ mit $\alpha_i < v_i < b_i < \beta_i$ für i=0(1)m+1. Die Anwendung des Lemmas mit $M_1 = Z_i$ und $M_3 = B_i$, i=0(1)m+1, garantiert die Existenz eines Elementes $f_m^i \in C^\infty[B_i]$ mit der Eigenschaft

$$1 = \|f_m^i\|_{B_i} = f_m^i(x) \text{ für } x \in Z_i \text{ und } 0 \leq f_m^i(x) \leq 1 \text{ sonst.}$$

Nun wird für $x \in [a,b]$ definiert

$$f_m^*(x) := \begin{cases} c_m \cdot \operatorname{sgn}\{q_t(x) - q_m(x)\} \cdot f_m^i(x) & \text{falls } x \in B_i, \\ & i \in \{0,1,\ldots,m+1\} \\ 0 & \text{sonst} \end{cases}$$

Offenbar ist auch $f_m^* \in C^\infty[a,b]$, ferner $\|f_m^*\|_{[a,b]} = c_m$ und $f_m^* \in A_{Q_m}^{-1}(0)$.

3. Konstruktion einer t-Alternierenden

Es seien die Größen $g, c_t^*, c_t, U(x_s)$ wie im Teil B3. vom Beweis zu Satz 4.1 gewählt.
In $U := U(x_s)$ wähle man Punkte r_i (i=1(1)t+1) mit

$$h_1 =: r_0 < r_1 < \ldots < r_{t+1} < r_{t+2} := h_2$$

Ferner sei

$U_i := [r_{i-1}, r_i]$, i=1(1)t+2, und $\tilde{U}_i := [\tilde{r}_{i-1}, \tilde{r}_i] \subset U_i$ mit

$r_{i-1} < \tilde{r}_{i-1} < \tilde{r}_i < r_i$ für i=1(1)t+2.
Die Anwendung des Lemmas mit $M_1 = U_i$ und $M_3 = \tilde{U}_i$, i=1(1)t+2, garantiert wieder die Existenz eines $f_t^i \in C^\infty[U_i]$ mit der Eigenschaft

$$1 = \|f_t^i\|_{U_i} = f_t^i(x) \text{ für } x \in \tilde{U}_i \text{ und } 0 \leq f_t^i(x) \leq 1 \text{ sonst.}$$

Nun wird für $x \in [a,b]$ definiert

$$f_t^*(x) := \begin{cases} (-1)^i \cdot c_t \cdot f_t^i(x) & \text{falls } x \in U_i,\ i \in \{1,2,\ldots,t+2\} \\ 0 & \text{sonst} \end{cases}$$

Auch hier gilt $f_t^* \in C^\infty[a,b]$, ferner $\|f_t^*\|_{[a,b]} = c_t$ und $f_t^* \in A_{Q_t}^{-1}(0)$.

4. Definition der gesuchten Funktion $f \in L(q_t, q_m)$.

$$\text{Sei } f_m(x) := \begin{cases} f_m^*(x) & \text{für } x \in [a,b]-U \\ f_t^*(x) + (q_t(x) - q_m(x)) & \text{für } x \in U \end{cases}$$

und $f_t(x) := f_m(x) - (q_t(x) - q_m(x))$ für $x \in [a,b]$.

Schließlich sei definiert $f(x) := f_m(x) + q_m(x)$ für $x \in [a,b]$.
Wie im Beweis von Satz 4.1 ergibt sich $f \in L(q_m) \cap L(q_t)$.
Jetzt gilt jedoch zusätzlich $f \in C^{\min(\ell,h)}[a,b]$; denn $(q_t - q_m) \in C^{\min(\ell,h)}$ und $f_m^* \in C^\infty$ und $f_t^* \in C^\infty$.

-- q.e.d. --

Bemerkung 4.2: Für den Spezialfall, daß q_t und q_m gewöhnliche algebraische Polynome sind, ergibt sich also, daß es immer möglich ist, eine Rivlin-Funktion f mit $f \in C^\infty[a,b]$ zu finden, falls $q_t - q_m$ sein Vorzeichen an mindestens m+1 Stellen in (a,b) wechselt. Dieses Ergebnis ist eine Verschärfung der Aussage von D.SPRECHER [1969].

In der Arbeit von DEUTSCH, MORRIS, SINGER [1969] klingt im Anschluß an das notwendige Kriterium von T.J.RIVLIN (vergl.

Bemerkung 3.2) die Frage an, inwieweit dieses Kriterium auch hinreichend sei. Für den Fall eines u-Polynompaares haben wir oben die Vermutung bestätigt. Daß die zitierte Vermutung bereits für die Vorgabe von Funktionentripeln nicht zutrifft, zeigten die Sätze 3.3 - 3.5, die außer dem Rivlin-Kriterium noch gewisse weitere "Verzahnungen" der Vorgabefunktionen als notwendig bewiesen. Es stellt sich nun heraus, daß die sich aus den Sätzen 3.3 - 3.5 ergebenden Zusatzforderungen auch hinreichend sind für die Existenz einer Rivlin-Funktion bei vorgegebenem u-Polynomtripel.

Zur Vorbereitung von Satz 4.3 (s.u.) untersuchen wir folgende Situation (V):

(V): Die Funktionen $u_\ell \in R$, $\ell=0(1)n \geq 2$, mögen ein erweitertes Haarsystem auf $[a,b]$ bilden. Es sei
$Q_\ell := \operatorname{span}(u_0,\ldots,u_\ell)$, $\ell=0(1)n$.
Ferner seien gegeben u-Polynome
$q_i = \sum_{s=0}^{i} a_s u_s \in Q_i$ mit $a_i \neq 0$. für $i=r,r+1,r+2$ und $0 \leq r,r+2 \leq n$.

Mit x_i ($x_k<x_m$ für $k<m$) seien die Nullstellen in $[a,b]$ von $d_{20}(x):=q_{r+2}(x)-q_r(x)$ bezeichnet, entsprechend mit y_i bzw. z_i die Nullstellen von $d_{10}(x):=q_{r+1}(x)-q_r(x)$ resp. $d_{21}(x):=q_{r+2}(x)-q_{r+1}(x)$.

Offenbar gilt $d_{21}(x)=d_{20}(x)-d_{10}(x)$.

Wir sagen nun, die gegebenen u-Polynome erfüllen die Nullstellenbedingung (NB), wenn die Situation (V) vorliegt und gilt:

(NB): 1) $Z(d_{21}) = r+2$, $Z(d_{10}) = r+1$, $Z(d_{20}) \geq r+1$
2) alle Nullstellen von $d_{21}(x)$ und $d_{10}(x)$ liegen im Innern von $[a,b]$
3) alle Nullstellen von $d_{20}(x)$ sind einfach und mindestens $r+1$ davon liegen im Innern von $[a,b]$

Bezüglich (NB)3. werden wir im Folgenden o.B.d.A. voraussetzen, daß gilt $x_1 > a$, und ferner lassen wir $x_{r+2} \geq b$ zu.

Das ist keine Einschränkung, da der Fall $x_1 \leq a$ bzgl. des Intervalls "symmetrisch" dazu liegt.

Gelte nun (V) und (NB) mit $x_1 > a$. Dann gibt es zu jeder Differenzfunktion

$$d_{k\ell}(x) = q_{r+k}(x) - q_{r+\ell}(x) \text{ mit } 0 \leq \ell < k \leq 2$$

innere Punkte $h_i \in [a,b]$, $i=1(1)m \geq \ell+1$, derart, daß $d_{k\ell}(x)$ in h_i sein Vorzeichen wechselt.
Es sei nun definiert:

$$J_1^{k\ell} := [a,h_1],\ J_{i+1}^{k\ell} := [h_i,h_{i+1}] \text{ für } i=1(1)\ell,\ J_{\ell+2}^{k\ell} := [h_{\ell+1},b]$$

(Im Falle $x_1=a$ wären nur die Indizes entsprechend zu modifizieren).

Ferner kürzen wir ab:

$$\overline{J}_i^{k\ell} := \{x \in J_i^{k\ell} \text{ und } d_{k\ell}(x) > 0\}$$

$$\underline{J}_i^{k\ell} := \{x \in J_i^{k\ell} \text{ und } d_{k\ell}(x) < 0\}$$

$$\overline{D}_i := \overline{J}_i^{10} \cap \overline{J}_i^{20},\ \underline{D}_i := \underline{J}_i^{10} \cap \underline{J}_i^{20},\ \text{jeweils i passend aus } \{1,\ldots,r+2\}$$

Mithilfe der soeben definierten Teilmengen des Intervalls [a,b] werden nun folgende Größen definiert:

$$n_1 := \min\{\min_i \|d_{10}\|_{\underline{D}_i},\ \min_s \|d_{10}\|_{\overline{D}_s}\}$$

$$n_2 := \min\{\min_i \|d_{20}\|_{\underline{D}_i},\ \min_s \|d_{20}\|_{\overline{D}_s}\}$$

$$n_3 := \min_{1 \leq i \leq r+3} \|d_{21}\|_{J_i^{21}}$$

Es bedeutet also z.B. n_1 das Minimum der T-Norm-Werte, die d_{10} annimmt auf allen Teilintervallen von [a,b], in denen $d_{10}(x)$ und $d_{20}(x)$ gleiches Vorzeichen aufweisen.

Bevor wir den Satz 4.3 angeben, werden noch zwei Bedingungen formuliert. Wir sagen, die in (V) gegebenen u-Polynome erfüllen die Bedingung der Relativen Nullstellenlage (NL), wenn gilt:

(NL): Für die Nullstellen von $d_{10}(x)$ und $d_{20}(x)$ gilt (NB) und

a) falls $Z(q_{r+2}-q_r)=r+1$, dann ist $\max(x_i,y_i)<\min(x_{i+1},y_{i+1})$ für $i=1(1)r$

b) falls $Z(q_{r+2}-q_r)=>r+1$, dann ist

entweder 1) $\max(x_i,y_i)<\min(x_{i+1},y_{i+1})$, $i=1(1)r$ und $x_{r+2}>\max(x_{r+1},y_{r+1})$

oder 2) $\max(x_{i+1},y_i)<\min(x_{i+2},y_{i+1})$, $i=1(1)r$ und $x_1<\min(x_2,y_1)$

Die in (V) gegebenen Funktionen erfüllen die Randbedingung (RB), wenn gilt

(RB): An genau einem Intervallendpunkt R_1 von $[a,b]$ haben $d_{10}(x)$, $d_{20}(x)$ und $d_{21}(x)$ gleiches Vorzeichen.

Es sei kurz darauf hingewiesen, daß (NB) dem Rivlin-Kriterium entspricht, ferner, daß der Nachweis von (NL) und (RB) bei erfülltem (NB) i.A. in der Praxis keine Zusatzschwierigkeiten mit sich bringt, wo-hingegen (NB) meistens nur numerisch zu verifizieren ist. Nur in einigen Spezialfällen kann das Erfülltsein von (NB) ohne explizite Nullstellenberechnung direkt z.B. aus den Polynomkoeffizienten erbracht werden. Für den Fall von algebraischen Polynomen sei hier etwa an den STURMschen Satz (siehe z.B. in Van der WAERDEN [1964], Seite 232) oder an die Regel von DESCARTES (siehe z.B. in KARLIN-STUDDEN [1966], Seite 27) erinnert. Regeln dieser Art wurden auch schon für Teilklassen von Haarsystemen bewiesen (siehe etwa in KARLIN-STUDDEN, Seite 25 ff).

Wenden wir uns nun der Formulierung des Satzes 3 zu:

SATZ 4.3: Es liege die Situation (V) vor und es mögen die vorstehenden Bezeichnungen gelten. Dann ist folgende Charakterisierung möglich.

Beh.: Es existiert ein $f \in R$ mit $f \in L(q_r,q_{r+1},q_{r+2})$ genau dann, wenn Folgendes erfüllt ist:

1. Die vorgegebenen u-Polynome erfüllen die Forderungen (NB),(NL),(RB).
2. Es lassen sich positive reelle Konstanten k_1 und k_3 angeben mit folgenden Eigenschaften:

   a) $k_1 \leq n_1$, $k_3 \leq n_3$ und $k_1 + k_3 =: k_2 \leq n_2$

   b) Es existiert ein nichtentartetes Teilintervall $U \subset [a,b]$ mit der Eigenschaft, daß gilt

   $|d_{20}(x)| \leq k_2$ und $|d_{21}(x)| \leq k_3$ für alle $x \in U$.

   c) Sei $\overline{E}_i^{21} := \{x \in J_i^{21} \text{ mit } k_3 \leq |d_{21}(x)|\}$

   und $\underline{E}_\ell^{21} := \{x \in \underline{J}_\ell^{21} \text{ mit } k_3 \leq |d_{21}(x)|\}$, ℓ und i passend aus $\{1,\ldots,r+3\}$

   Dann gilt:

   für jedes i ist $\{x \in \overline{E}_i^{21} \text{ mit } d_{10}(x) \leq k_1\} \neq \phi$ und

   für jedes ℓ ist $\{x \in \underline{E}_\ell^{21} \text{ mit } -d_{10}(x) \leq k_1\} \neq \phi$

Geometrische Bedeutung der Bedingungen

1. (NB) und (RB) stimmen die "Breite" des Approximationsintervalls und die vorgegebenen u-Polynome aufeinander ab.
2. (NL) und (2b) fordern eine bestimmte "Verzahnung" der Nullstellen der Differenzfunktionen. Dabei besorgt (NL) eine gewisse Trennbarkeit und (2b) eine Verdichtung der Nullstellen an mindestens einer Stelle in [a,b], denn das in (2b) geforderte U wird i.A. eine "Nachbarschaft" eines Nullstellenpaares (x,z) von Nullstellen von $d_{20}(x)$ und $d_{21}(x)$ sein.
3. (2a) und insbesondere (2c) korreliert die Verläufe der Graphen der Differenzfunktionen (und damit der gegebenen u-Polynome) außerhalb der Nullstellen (resp. Schnittstellen) und sorgt dafür, daß sich die Funktionen nur "zulässig weit" voneinander "entfernen". So wird z.B. in (2c) gefordert, daß dort, wo $d_{21}(x)$ den Wert k_3 überschreitet, jeweils mindestens an einer Stelle dieser Teilintervalle zugleich $d_{10}(x)$ nicht größer als k_1 ist.

Beweis von Satz 4.3:

A) Notwendigkeit: Voraussetzung ist, daß ein $f \in L(q_r, q_{r+1}, q_{r+2})$ existiert.

zu 1.: (NB) folgt aus Satz 3.2
(NL) folgt aus Satz 3.4
(RB) folgt aus Satz 3.5

zu 2.: Diese Bedingungsgruppe folgt unmittelbar aus Satz 3.3.
Man betrachte nämlich die dortigen Aussagen für den Fall t=2 , setze $k_1 := c_{r,r+1}$, $k_2 := c_{r,r+2}$, $k_3 := c_{r+1,r+2}$ und beachte die Definitionen der Konstanten n_1, n_2, n_3.

B) Hinlänglichkeit:

Hier spielt die Äquivalenzaussage von Satz 1.5 eine wesentliche Rolle. Der Beweis wird in zwei Schritten geführt:

1. Es wird zunächst gezeigt, daß aufgrund der Satzvoraussetzungen je genau k+2 "qualifizierte" k-Abweichungsintervalle S_ℓ^k (ℓ=1(1)k+2) k=r(1)r+2 für die gemäß Satz 1.5 zu konstruierenden k-Nullalternierenden f_k existieren. Dabei sind unter "qualifizierten" k-Abweichungsintervallen solche Teilintervalle von [a,b] zu verstehen, in denen in Kompatibilität zu Satz 3.3 (vergleiche dortigen Beweis) die Abweichungspunkte der zu konstruierenden k-Alternierenden f_k gewählt werden können.
2. Die k-Alternierenden f_k, k=r(1)r+2, werden konstruiert und daraus eine gesuchte Funktion $f \in L(q_r, q_{r+1}, q_{r+2})$ bestimmt.

Wir setzen im folgenden Beweis o.B.d.A. voraus, daß für die erste Nullstelle x_1 von $d_{20}(x)$ im Intervall [a,b] gilt $x_1 > 0$. Das bedeutet keine Einschränkung der Allgemeinheit des Beweises, da der Fall x_1=a durch einfache Indexmodifikationen sich erledigen läßt. Ebenso kann o.B.d.A. bezüglich (RB) angenommen werden, daß gilt R_1=a.(vgl.auch Bemerkung 3. zu Satz 3.4).

zu 1.):

a) Die Nullstellen x_i, $i=1(1)Z(d_{20})\geq r+1$, und y_ℓ, $\ell=1(1)r+1$, sind gemäß Bedingung (NB) Punkte aus dem halboffenen Intervall $(a,b]$. Wir definieren folgende Teilintervalle:

$D_1 := [a,\min(x_1,y_1))$

$D_i := (\max(x_{i-1},y_{i-1}),\min(x_i,y_i))$ für $i=2(1)r+1$

$$D_{r+2} := \begin{cases} (\max(x_{r+1},y_{r+1}),b] & \text{falls } Z(d_{20})=r+1 \\ (\max(x_{r+1},y_{r+1}),b) & \text{falls } Z(d_{20})=r+2 \text{ und } d_{20}(b)=0 \\ (\max(x_{r+1},y_{r+1}),x_{r+2}) & \text{sonst} \end{cases}$$

Aus (NB) und (NL) zusammen folgt nun

(a1) $D_\ell \neq \phi$ für $\ell=1(1)r+2$

(a2) aus $x \in D_i$, $y \in D_\ell$, $i<\ell$ folgt $x<y$

(a3) die Funktionen $d_{20}(x)$ und $d_{10}(x)$ haben auf D_ℓ, $\ell=1(1)r+2$, jeweils konstantes Vorzeichen.

Wegen (RB) und der Voraussetzung $R_1=a$ folgt sogar, daß gilt

$$\operatorname{sgn} d_{20}(x)=\operatorname{sgn} d_{10}(x) \text{ für } x \in D_\ell,\ \ell=1(1)r+2.$$

Mit o.B.d.A.: $\operatorname{sgn} d_{10}(R_1)=-1$ (das entspricht dem Grundtyp I)

gilt dann:

$$D_i=\overline{D}_i=\overline{J}_i^{10} \cap \overline{J}_i^{20} \neq \phi \quad \text{für } i = \begin{cases} 2(2)r+1 & \text{falls } r \text{ ungerade} \\ 2(2)r+2 & \text{falls } r \text{ gerade} \end{cases} \qquad (+)$$

$$D_i=\underline{D}_i=\underline{J}_i^{10} \cap \underline{J}_i^{20} \neq \phi \quad \text{für } i = \begin{cases} 1(2)r+2 & \text{falls } r \text{ ungerade} \\ 1(2)r+1 & \text{falls } r \text{ gerade} \end{cases} \qquad (++)$$

Weil nach Satzvoraussetzung $k_1 \leq n_1$, $k_2 \leq n_2$ und $k_1<k_2$ gilt, folgt daraus, daß auch gilt:

$\overline{S}_i^r := \{x \in D_i;\ k_1 \leq |d_{10}(x)| \text{ und } k_2 \leq |d_{20}(x)|\} \neq \phi$ für alle Werte i aus (+)

$\underline{S}_i^r := \{x \in D_i;\ k_1 \leq |d_{10}(x)| \text{ und } k_2 \leq |d_{20}(x)|\} \neq \phi$ für alle Werte i aus (++)

Insgesamt also existieren genau r+2 nichtleere qualifizierte disjunkte Intervalle, die den Forderungen an die r-Ab-

weichungsintervalle $S^r(f_r)$ aus Satz 3.3 genügen, falls die zu konstruierenden k-Alternierenden f_k, k=r(1)r+2, so gewählt werden, daß die Beziehungen $k_1=\|f_r\|-\|f_{r+1}\|$, $k_2=\|f_r\|-\|f_{r+2}\|$ erfüllt sind.

b):Die Satzvoraussetzungen (NB) bezüglich $d_{21}(x)$, zusammen mit der Wahl von k_3 und der Voraussetzung (2a) ergeben unmittelbar die Existenz von genau r+3 nichtleeren disjunkten qualifizierten Intervallen, die den Anforderungen an die (r+1)-Abweichungsintervalle $S^{r+1}(f_{r+1})$ aus Satz 3.3 mit $k_3=\|f_{r+1}\|-\|f_{r+2}\|$ genügen, nämlich

$$\overline{S}_i^{r+1}:=\{x\in\overline{E}_i^{21} \text{ mit } d_{10}(x)\leq k_1\} \text{ für } i=\begin{cases}2(2)r+3 \text{ falls r ungerad}\\ 2(2)r+2 \text{ falls r gerade}\end{cases}$$

$$\underline{S}_i^{r+1}:=\{x\in\underline{E}_i^{21} \text{ mit } -d_{10}(x)\leq k_1\} \text{ für } i=\begin{cases}1(2)r+2 \text{ falls r ungerad}\\ 1(2)r+3 \text{ falls r gerade}\end{cases}$$

c):Die Satzvoraussetzung (2b) besagt, daß in $U=:[a_1,b_1]$ jeder Punkt qualifiziert ist im Sinne Satz 3.3, um zu einem der r+4 Abweichungsintervalle $S_i^{r+2}(f_{r+2})$ zu gehören. Man wähle etwa eine Folge von inneren Punkten h_i, i=1(1)r+4, mit $h_1<h_2<\ldots<h_{r+4}$ aus U und setze $S_i^{r+2}:=\{h_i\}$, i=1(1)r+4. —

Hinweis: Im Paragraphen 5 werden für drei Tschebyscheffpolynome 2. Art die "qualifizierten" Intervalle als Beispiel explizit angegeben mit Übersichtsskizze.

zu 2.):

Man wähle

$$g_2>2\max\{\|d_{10}\|,\|d_{20}\|,\|d_{21}\|\},$$

wobei die Norm bzgl. des Intervalls [a,b] gerechnet ist, und setze

$$g_1:=g_2+k_3, \quad g_0:=g_2+k_2=g_1+k_1,$$

wobei die k_i (i=1(1)3) die Konstanten aus der Satzvoraussetzung sind.

Dann gilt $g_2 < g_1 < g_0$.
Es werden im Folgenden die i-Alternierenden f_i, i=r(1)r+2, so konstruiert, daß gilt $\|f_{r+i}\| = g_i$.
Gemäß Teil 1. existieren genau i+2 qualifizierte Intervalle S_ℓ^i, ℓ=1(1)i+2, i=r(1)r+2. Man wähle nun Punkte $s_\ell^i \in S_\ell^i$, ℓ=1(1)i+2, für i=r und i=r+1 mit der Eigenschaft $s_\ell^r < s_{\ell+1}^{r+1}$, ℓ=1(1)r+2. Ferner sei $s_\ell^{r+2} := h_\ell$, ℓ=1(1)r+4, mit h_ℓ gemäß Teil 1.c).

In Anwendung von Satz 1.5 und Satz 2.3 ist nun eine stetige Funktion $f_r(x)$ auf [a,b] zu konstruieren, welche die folgenden Eigenschaften (E) hat:

(E): a) Es gilt $f_r(s_\ell^r)=(-1)^\ell g_0$, ℓ=1(1)r+2, ansonsten $|f_r(x)| \le g_0$ für $x \in [a,b]$
b) Es gilt $f_r(s_\ell^{r+1})-d_{10}(s_\ell^{r+1})=(-1)^\ell g_1$, ℓ=1(1)r+3, ansonsten $|f_r(x)-d_{10}(x)| \le g_1$ für $x \in [a,b]$
c) Es gilt $f_r(s_\ell^{r+2})-d_{20}(s_\ell^{r+2})=(-1)^\ell g_2$, ℓ=1(1)r+4, ansonsten $|f_r(x)-d_{20}(x)| \le g_2$ für $x \in [a,b]$

Die Funktionen $d_{21}(x)$, $d_{20}(x)$, $d_{10}(x)$ sind als differenzierbare Funktionen über [a,b] Lipschitz-beschränkt mit Lipschitz-Konstanten L_3, L_2 resp. L_1, die nach unten beschränkt sind durch die T-Normen der Ableitungen der entsprechenden Funktionen, also $\|d'_{\ell k}\| \le L_i$, i=k+ℓ, $0 \le k < \ell \le 2$. Es sei

$$L_m := \max\{\|d'_{10}\|, \|d'_{20}\|, \|d'_{21}\|\} .$$

Man wähle nun $L > L_m$ und bestimme die Konstanten

$$\varepsilon_i := g_i / (2L) \text{ für } i=0 \text{ und } i=1.$$

Für $x \in [a,b]$ und $i \in \{0,1\}$ und ℓ=1(1)r+2+i wird nun definiert

$$\tilde{f}_{r+i}(x) := \begin{cases} (-1)^\ell g_i \ (x-s_\ell^{r+i}+\varepsilon_i)/\varepsilon_i & \text{für } x \in [s_\ell^{r+i}-\varepsilon_i, s_\ell^{r+i}] \\ (-1)^{\ell+1} g_i (x-s_\ell^{r+i}-\varepsilon_i)/\varepsilon_i & \text{für } x \in [s_\ell^{r+i}, s_\ell^{r+i}+\varepsilon_i] \\ 0 & \text{sonst} \end{cases}$$

Dabei ist natürlich $s_1^{r+i}-\varepsilon_i$ als $\max(a, s_1^{r+i}-\varepsilon_i)$ und $s_{r+2+i}^{r+i}+\varepsilon_i$ als $\min(b, s_{r+2+i}^{r+i}+\varepsilon_i)$ aufzufassen.

Für $i \in \{0,1\}$ gilt dann: es ist $\tilde{f}_{r+i} \in C[a,b]$ und es ist

$$\tilde{f}_{r+i}(s_\ell^{r+i}) = (-1)^\ell g_i, \quad \ell=1(1)r+2+i, \text{ ferner}$$

$$|\tilde{f}_{r+i}(x)| \leq g_i \text{ für } x \in [a,b]$$

Somit ist $\|\tilde{f}_{r+i}\| = g_i$ und $\tilde{f}_{r+i}$ ist eine $(r+i)$-Nullalternierende.

Weiter sei definiert für $x \in [a,b]$:

$$\tilde{f}_{r+2}(x) := \begin{cases} g_2(x-a_1)/(h_1-a_1) & \text{für } x \in [a_1,h_1] \\ g_2(h_{2\ell-1}+h_{2\ell}-2x)/(h_{2\ell}-h_{2\ell-1}) & \text{für } x \in [h_{2\ell-1},h_{2\ell}] \text{ und } \ell=1(1)\,[(r+4)/2] \\ g_2(2x-h_{2\ell}-h_{2\ell+1})/(h_{2\ell+1}-h_{2\ell}) & \text{für } x \in [h_{2\ell},h_{2\ell+1}] \text{ und } \ell=1(1)\,[(r+3)/2] \\ (-1)^{r+1} g_2(b_1-x)/(b_1-h_{r+4}) & \text{für } x \in [h_{r+4},b_1] \\ 0 & \text{sonst} \end{cases}$$

Wie oben ergibt sich: $\tilde{f}_{r+2}$ ist eine stetige $(r+2)$-Alternierende mit dem T-Normwert g_2.

Bevor die endgültigen Alternierenden definiert werden, die dem Satz 1.5 entsprechen, müssen noch einige Hilfsfunktionen erklärt werden.

Zunächst sei gesetzt: $\bar{f}_{r+1}(x) := \begin{cases} \tilde{f}_{r+2}(x)+d_{21}(x) & \text{für } x \in U \\ \tilde{f}_{r+1}(x) & \text{sonst in } [a,b] \end{cases}$

Es ist $\|d_{21}\|_U \leq g_1-g_2$ nach Wahl von U. Deshalb gilt auch

$$\|\bar{f}_{r+1}\| = g_1 \text{ und } \bar{f}_{r+1}(s_\ell^{r+1}) = (-1)^\ell g_1 \text{ für } \ell=1(1)r+3. \quad (+++)$$

Wegen der Lipschitzbeschränktheit der auftretenden Differenzpolynome läßt sich nun nötigenfalls die Funktion $\bar{f}_{r+1}(x)$ in hinreichend kleinen Umgebungen ihrer höchstens zwei möglichen Unstetigkeitsstellen a_1 und b_1 so in eine stetige Funktion $f^*_{r+1}(x)$ abändern, daß insbesondere die Eigenschaften

(+++) auch für $f^*_{r+1}(x)$ statt $\bar{f}_{r+1}(x)$ gelten.

Mit $B^i_\ell := [s^{r+i}_\ell - \varepsilon_i,\ s^{r+i}_\ell + \varepsilon_i]$, $\ell=1(1)r+2+i$, $i \in \{0,1\}$ wird gesetzt

$$f^*_r(x) := \begin{cases} f^*_{r+1}(x)+d_{10}(x) & \text{für } x \in B_\ell := B^1_\ell - B^1_\ell B^o_\ell,\ \ell=1(1)r+2 \\ \tilde{f}_r(x) & \text{sonst für } x \in [a,b] \end{cases}$$

Wegen der Wahl von $s^{r+1}_\ell \in S^{r+1}_\ell$, dessen Intervalleigenschaften und dem Verhalten von $f^*_{r+1}(x)$ auf B^1_ℓ, $\ell=1(1)r+3$, ist sichergestellt, daß gilt:

$$\|f^*_r\| = g_o \text{ und } f^*_r(s^r_\ell) = (-)^\ell g_o,\ \ell=1(1)r+2 \qquad (*)$$

Mit der gleichen Begründung wie oben kann $f^*_r(x)$ in genügend kleinen Umgebungen der höchstens $2r+4$ möglichen Unstetigkeitsstellen (nämlich den Endpunkten der B_ℓ) übergeführt werden in eine stetige Version $f_r(x)$, ohne daß die Eigenschaften (*) für $f_r(x)$ statt $f^*_r(x)$ verloren gehen.

Mithilfe von $f_r(x)$ werden nun auch die beiden anderen Alternierenden f_{r+1} und f_{r+2} definiert auf $[a,b]$ gemäß

$$\begin{aligned} f_{r+1}(x) &:= f_r(x) - d_{10}(x) = f_r(x) - (q_{r+1}(x) - q_r(x)) \\ f_{r+2}(x) &:= f_{r+1}(x) - d_{21}(x) = f_{r+1}(x) - (q_{r+2}(x) - q_{r+1}(x)) \\ &= f_r(x) - (q_{r+2}(x) - q_r(x)) \end{aligned}$$

Gemäß der vorstehenden Konstruktionen und den Eigenschaften der Differenzpolynome in den Intervallen U und S^{r+i}_ℓ, $\ell=1(1)r+i+2$ und $i \in \{0,1\}$, folgt nun unmittelbar, daß f_{r+i}, $i=0(1)2$, den Eigenschaften (E) genügen. Der Alternantensatz besagt dann $f_\ell \in A^{-1}_{Q_\ell}(0)$, $\ell=r(1)r+2$.

Weiterhin gilt nach Konstruktion $f_i = f_\ell - (q_i - q_\ell)$, $r \le \ell < i \le r+2$, und gemäß der Wahl von g_o, g_1, g_2 wegen $k_1 = \|f_r\| - \|f_{r+1}\|$, $k_2 = \|f_r\| - \|f_{r+2}\|$ sowie $k_3 = \|f_{r+1}\| - \|f_{r+2}\|$ auch $\|f_r\| > \|f_{r+1}\| > \|f_{r+2}\|$.

Definieren wir schließlich $f(x) := f_r(x) + q_r(x)$, $x \in [a,b]$, so ist $f \in C[a,b]$ und wegen $f_i + q_i = f_\ell + q_\ell$, $r \le \ell < i \le r+2$ und den Eigen-

schaften von f_i, i=r(1)r+2 gilt $f \in L(q_r, q_{r+1}, q_{r+2})$.

-- q.e.d. --

Bemerkung 4.3: Die gemäß Satz 4.3 existierende Rivlin-Funktion f ist nicht eindeutig bestimmt.

Zum Abschluß dieses Paragraphen soll noch ein weiterer Charakterisierungssatz angegeben werden, den SPRECHER [1969a] bewies und der als Spezialfall in Satz 4.3 enthalten ist, nämlich für den Fall, daß

(1) $u_i(x) := x^i$, i=0(1)n, (2) r=0 und (3) [a,b]=[0,1].

Wir zitieren den Satz in der Formulierung von D.SPRECHER [1969a] und verweisen auch auf den dortigen Beweis:

SATZ 4.4: Sei R=C[a,b] mit [a,b] = [0,1].
Sei $p_o(x)=a_o$, $p_1(x)=b_o+b_1x$, $p_2(x)=c_o+c_1x+c_2x^2$ mit reellen Koeffizienten und sei $b_1 \neq 0$, $c_2 \neq 0$.

Beh.: Es gilt $L(p_o,p_1,p_2) \neq \phi$ genau dann, wenn

a) es Punkte gibt $0<z_1<z_2<1$, sodaß
$p_1(z_\ell)=p_2(z_\ell)$, (ℓ=1,2), und
b) es gilt:
$z_1/2<(a_o-b_o)/b_1<(z_1+z_2)/2$ falls ($b_1>0, c_2<0$ oder $b_1<0, c_2>0$), oder $(z_1+z_2)/2<(a_o-b_o)/b_1<(z_1+1)/2$ falls ($b_1<0, c_2<0$ oder $b_1>0, c_2>0$)

Wie der Bedingungskomplex des Satzes 4.3 aus den Bedingungen des obigen Satzes 4.4 für den vorliegenden Spezialfall r=0, $u_i(x)=x^i$ mit i=0(1)2, hervorgeht, soll nun andeutungsweise skizziert werden. Dabei nehmen wir o.B.d.A. an: $a_o=0$ und es gelte in Satz 4.4 die Bedingung b) mit $b_1>0$, $c_2<0$.

1.: Gelten die Bedingungen von Satz 4.4, so auch die von Satz 4.3:

Die Monome $u_i(x)=x^i$, i=0(1)2, bilden ein ECT-System.

zu (RB): Wegen $0<z_1/2<-b_0/b_1<(z_1+z_2)/2$, $b_1>0$, $c_2<0$ ist $b_0<0$ und wegen $z_1>0$ ist $p_2(x)<p_1(x)$ für $x<z_1$, somit insgesamt sgn $d_{21}(0)$=sgn $d_{10}(0)$=sgn $d_{20}(0) = -1$, ferner sgn $d_{21}(b) = -$sgn $d_{10}(b)$.
Somit gilt $R_1 = a$ und (RB) ist erfüllt.

zu (NB): Die zwei Nullstellen von $d_{21}(x)$ werden in Bedingung Satz 4.4a) explizit gefordert. Die Nullstelle y_1 von $d_{10}(x)$ wird ebenfalls im Innern von [0,1] garantiert durch $z_1/2<y_1 = -b_0/b_1<(z_1+z_2)/2<z_2$. Schließlich ist $d_{20}(z_2) = d_{10}(z_2)$, ferner $d_{10}(z_2)>0$ und $d_{10}(0)<0$. Folglich existiert eine Nullstelle x_1 von $d_{20}(x)$ mit $0<x_1<z_2<1$. Insgesamt also gilt (NB).

zu (NL): Es ist $d_{20}(z_2)>0$ und $c_2<0$. Also gilt für eine zweite Nullstelle x_2 von $d_{20}(x)$ die Ungleichung $z_2<x_2$. Da auch $y_1<x_2$ gilt, folgt insgesamt $\max(x_1,y_1)<x_2$, somit gilt (NL).

zur 2. Bedingungsgruppe von Satz 4.3: SPRECHERS Bedingung

$$0 < z_1/2 < -b_0/b_1 < (z_1+z_2)/2$$

besagt gerade, daß gilt

$$|d_{10}(z_1)| < \min\{|d_{10}(0)|, d_{10}(z_2)\}.$$

Man wähle die Konstante k_1 des Satzes 4.4 wie folgt:

$$|d_{10}(z_1)| < k_1 < \min\{|d_{10}(0)|, d_{10}(z_2)\}$$

Es ist wegen $d_{10}(0) < d_{10}(z_2) < d_{10}(1)$ auch $k_1 \leq n_1$ erfüllt. Wegen der Stetigkeit und strengen Monotonie von $d_{10}(x)$ existieren Punkte t_1,t_2 sodaß $0<t_1<z_1<t_2<z_2$ und $|d_{10}(t_i)| = k_1$, (i=1,2), gilt. Sei nun eine Hilfsgröße m definiert durch $d_{21}(m) = \|d_{21}\|_{[z_1,z_2]}$. Es ist $z_1<m<z_2$.

Nun soll die Konstante k_3 gemäß Satz 4.4 bestimmt werden:

Fall 1: $y_1 = -b_0/b_1 = z_1 = x_1$:

Wähle α,β mit $t_1 < \alpha < z_1 < \beta < \min(t_2,m)$, sodaß gilt $\max\{|d_{21}(\alpha)|, d_{21}(\beta)\} \leq \min\{|d_{21}(0)|, |d_{21}(1)|\}$, was sich jedoch wegen der Stetigkeit von $d_{21}(x)$ erreichen läßt. Setze $k_3 := \max\{|d_{21}(\alpha)|, d_{21}(\beta)\}$.

Fall 2: $y_1 > z_1$, d.h. es gilt $z_1 < x_1 < y_1 < t_2$.

Wähle α,β mit $z_1 = \alpha < \beta < y_1$ derart, daß gilt $d_{21}(\beta) \leq \min\{|d_{21}(0)|, |d_{21}(1)|\}$ und setze $k_3 := d_{21}(\beta)$. (Vergl. BILD 4)

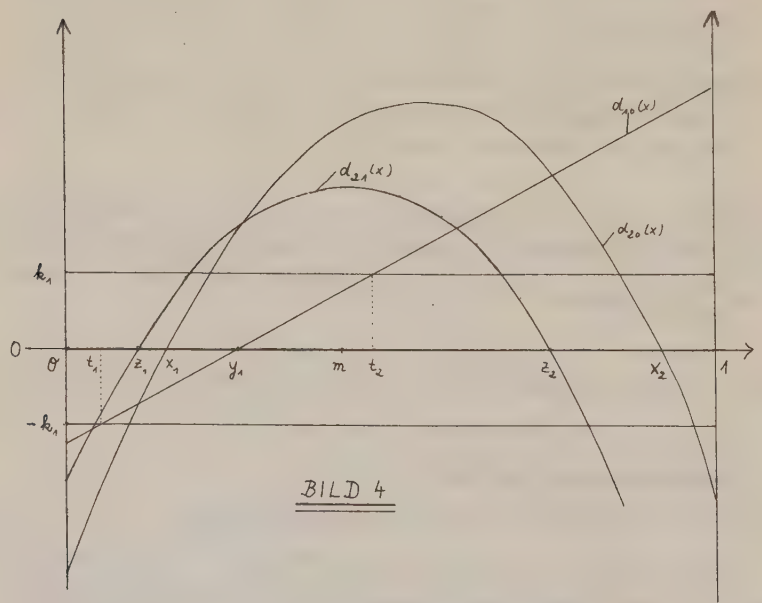

BILD 4

Fall 3: $y_1 < z_1$, d.h. es gilt $t_1 < y_1 < x_1 < z_1$.
Wähle α,β mit $y_1 < \alpha < \beta = z_1$ derart, daß gilt $|d_{21}(\alpha)| \leq \min\{d_{21}(m), |d_{21}(1)|\}$ und setze $k_3 := |d_{21}(\alpha)|$.

In allen drei Fällen gilt dann $k_3 \leq n_3$. Definiert man nun $U := [\alpha,\beta]$, so ergibt sich aus der Wahl der Konstanten k_1, k_3 des Intervalls U und der Tatsache $d_{20} = d_{21} + d_{10}$ durch

elementare Rechnung das Erfülltsein der restlichen Bedingungen der Bedingungsgruppe 2 von Satz 4.3.

2.: Sind die Bedingungen von Satz 44 verletzt, so auch die von Satz 43:

zu a: Ist Bedingung a) aus Satz 44 nicht erfüllt, so auch nicht (NB) in Satz 43.

zu b: Gelte in Satz 44 mit $b_1 > 0$, $c_2 < 0$ die Bedingung a), aber statt der Bedingung b) gelte $z_1/2 \geq (a_o - b_o)/b_1$: Die Polynome $p_0(x) \equiv 0$, $p_1(x) = x + 1/4$, $p_2(x) = (-20\,x^2 + 23\,x - 3)/3$ erfüllen diese Annahmen. Aber mit diesen Funktionen ist die Bedingung (RB) in Satz 43 nicht erfüllt.

Somit ergibt sich insgesamt, daß der Satz von D.SPRECHER [1969a] (obiger Satz 4.4) als Spezialfall in unserem Satz 4.3 enthalten ist.

§ 5 Beispiele

Vorgegeben sei wieder der Raum $R = C[a,b]$ - versehen mit der Tschebyscheff-Norm - , ferner lineare Teilräume $Q_\ell := \text{span}\,(u_0,u_1,\ldots,u_\ell)$ und $Q^*_{2\ell+1} := \text{span}\,(u^*_0,\ldots,u^*_{2\ell})$, $\ell \in \mathbb{N}_0$ mit $u_i(x) := x^i$ für $i=0(1)\ell$, sowie $u^*_0(x) := 1$ und $u^*_{2m-1}(x) := \cos mx$, $u^*_{2m}(x) := \sin mx$ für $m=1(1)\ell$.
Wir fragen nun, inwieweit bestimmte in den vorangegangenen Paragraphen entwickelte Kriterien für die Existenz einer Rivlin-Funktion $f \in R$ erfüllt sind, wenn als "vorgeschriebene Minimallösungen" $q_n \in Q_n$ (bzw. $q^*_n \in Q^*_n$) für f aus Q_n (resp. Q^*_n) die Repräsentanten einiger bekannter Funktionensysteme gewählt werden, die u.a. in der Approximationstheorie häufig Verwendung finden. (In den bisherigen Arbeiten von D.SPRECHER [1969;1969a] und DEUTSCH, MORRIS, SINGER [1969] befinden sich keine Untersuchungen dieser Art.)

Im einzelnen diskutieren wir die Monome $q_n(x)=x^n$, das Tripel (1, cos t, sin t), die Tschebyscheff-Polynome erster Art $q_n(x)=T_n(x)$ und zweiter Art $q_n(x)=U_n(x)$, sowie die Legendre-Polynome $q_n(x)=L_n(x)$.

Dabei wird sich zeigen, daß unter diesen Funktionensystemen die Tschebyscheff-Polynome zweiter Art hinsichtlich unserer Fragestellung eine besondere Rolle spielen.

5.1 Die Monome $q_n(x)=x^n$, $n \geq 0$, im Intervall $[a,b]=[0,1]$

Behauptung: Es gilt $\bigcap_{i=k}^{k+\ell} L(q_i) = \phi$, $k \geq 0$, $\ell \geq 1$

Beweis: Gemäß Satz 3.2 ist notwendig für die Existenz einer Rivlin-Funktion, daß jede Differenz x^n-x^m mit $k \leq m < n \leq k+\ell$ mindestens m+1 Vorzeichenwechsel im Innern von [a,b] hat. Die Funktion x^n-x^{n-1}, $n \geq 1$, hat höchstens n Nullstellen. Im Im Punkt x=1 haben alle Polynome x^n-x^{n-1} ($n \in \mathbb{N}$) eine Nullstelle. Somit kann x^n-x^{n-1} höchstens n-1 Nullstellen bzw.

Vorzeichenwechsel im Innern des Intervalls [0,1] besitzen.

5.2 Legendre-Polynome und Tschebyscheff-Polynome erster Art, [a,b] = [-1,+1]

$$L_n(x) := \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2-1)^n \quad , n \in \mathbb{N}_0, \text{ (Legendre-Polynome)}$$

$$T_n(x) := \cos (n \arccos x) \quad , n \in \mathbb{N}_0, \text{ (T-Polynome 1. Art)}$$

Behauptung: Sowohl für $q_n(x)=T_n(x)$ wie auch für $q_n(x)=L_n(x)$, $n \geq 0$, gilt $L(q_n) \cap L(q_{n+1}) = \phi$, d.h. es existiert keine Rivlin-Funktion.

Beweis: Es ist grad T_n = grad L_n = n ferner $T_n(1) = L_n(1)=1$ für $n \geq 0$. Somit ist das Kriterium des Satzes 3.2 verletzt.

5.3 Die Funktionen 1, cos t, sin t

Die Funktionen $u_0^*(t)=1$, $u_1^*(t)=\cos t$, $u_2^*(t)=\sin t$ bilden ein ECT-System in jedem Intervall [a,b], welches kein ganzzahliges Vielfaches von π oder die Null enthält, weil (vergl. Beispiel 6 von Abschnitt 2.3) gilt

$$W(u_0^*(t)) = 1 \quad \text{für alle } t \in [a,b]$$

$$W(u_0^*(t),u_1^*(t)) = \begin{vmatrix} 1 & \cos t \\ 0 & -\sin t \end{vmatrix} = -\sin t \neq 0 \text{ für alle } t \in [a,b]$$

$$W(u_0^*(t),u_1^*(t),u_2^*(t)) = \begin{vmatrix} 1 & \cos t & \sin t \\ 0 & -\sin t & \cos t \\ 0 & -\cos t & -\sin t \end{vmatrix} = \sin^2 t + \cos^2 t = 1 \quad \text{für alle } t \in [a,b]$$

Es existiert aber keine Rivlin-Funktion $f \in L(u_0^*,u_1^*,u_2^*)$, da wegen $u_0^*(t) \geq \max \{u_1^*(t),u_2^*(t)\}$ für alle $t \in [a,b]$ die Bedingung (NL) aus Satz 4.3 nie erfüllt sein kann. ———

Jedoch existiert z.B. im Intervall $[0,2\pi-0.01]$ ein $f \in L(u_1^*,u_2^*)$ aufgrund von Satz 4.1, da Punkte t_1,t_2 mit $0<t_1<t_2<2\pi-0.01$

existieren, sodaß $\cos t_\ell = \sin t_\ell, \ell \in \{1,2\}$, und weil $\{1,\cos t, \sin t\}$ ein T-System bilden für $0 \le t < 2\pi$.

5.4 Die Tschebyscheff-Polynome 2. Art, [a,b] = [-1,+1]

Die Tschebyscheff-Polynome zweiter Art sind auf [a,b] wie folgt definiert:

(5.1) $U_n(x) := [\sin((n+1)\arccos x)] \,/\, \sin \arccos x,\ n \in \mathbb{N}_0$

So ist z.B. $U_0(x) = 1$, $U_1(x) = 2x$, $U_2(x) = 4x^2-1$,
$U_3(x) = 8x^3-4x$

Behauptung 1: Für jedes Polynompaar (U_n, U_{n+2}), $n \ge 0$, existiert eine Funktion $f \in C^\infty[a,b]$ mit $f \in L(U_n, U_{n+2})$.

Beweis: Es gilt die Funktionalgleichung $U_{n+2}(x) - U_n(x) = 2 \cdot T_{n+2}(x)$, $n \ge 0$ (vergl. etwa ABRAMOWITZ-STEGUN [1968], Seite 778). Die Funktion $2T_{n+2}(x)$ verschwindet an den gleichen Stellen wie $T_{n+2}(x)$, sie hat also im Innern von [-1,+1] die Nullstellen

(5.2) $\qquad x_i = -\cos[(2i-1)\pi \,/\, (2n+4)],\ 1 \le i \le n+2.$

Ferner gilt $2T_{n+2}\left[\cos\left(i\pi/(n+2)\right)\right] = (-1)^i 2,\ 1 \le i \le n+2.$
Somit hat $U_{n+2}(x) - U_n(x)$ im Innern von [-1,+1] nicht weniger als n+1 Vorzeichenwechsel. Satz 4.2 liefert nun die Behauptung.

-- q.e.d. --

Behauptung 2: Für jedes $n \ge 0$ gilt $L(U_n, U_{n+1}) \neq \phi$

Beweis: Die Polynome $U_n(x)$, $n \ge 1$, haben die Nullstellen

(5.3) $\qquad x_i^{(n)} = -\cos(i/(n+1)\pi),\ 1 \le i \le n,$

wie sich sofort aus (5.1) ergibt, d.h. alle Nullstellen sind reell. Im Intervall $[0,\pi]$ ist die Funktion cos(x) streng monoton fallend, und wegen

$$0 < i/(n+1) < (i+1)/(n+1) < 1 \text{ für } 1 \le i \le n-1$$

gilt somit

(5.4) $x_{i+1}^{(n)} > x_i^{(n)}$, ferner $x_1^{(n)} > -1$ und $x_n^{(n)} < +1$, $i=1(1)n-1$.

Alle Nullstellen von $U_n(x)$, $n \geq 1$, sind somit einfach und liegen im Innern des Intervalls $[-1,+1]$.

Sei nun $n \in \mathbb{N}$ fest vorgegeben und $m := n+1$. Wegen

$$i/(m+1) < i/(n+1) < (i+1)/(m+1) \text{ für } 1 \leq i \leq n$$

gilt gemäß (5.3), daß

$$(5.5) \qquad x_1^{(m)} < x_1^{(n)} < x_2^{(m)} < \ldots < x_n^{(n)} < x_m^{(m)}.$$

Ferner gilt

$$(5.6) \qquad U_\ell(1) = \ell+1, \quad \ell \in \mathbb{N}_o ,$$

$$(5.7) \qquad U_\ell(-x) = (-1)^\ell U_\ell(x), \quad \ell \in \mathbb{N}_o .$$

Wegen (5.6) und (5.7) gilt also

$$(5.8) \qquad \operatorname{sgn} U_n(-1) = -\operatorname{sgn} U_m(-1)$$

und, weil alle Nullstellen der Polynome $U_\ell(x)$, $\ell \in \mathbb{N}$, gemäß (5.3) im Innern von $[-1,+1]$ liegen, folgt also, daß $|U_\ell(x)|$ im Intervall $[-1, x_1^{(\ell)}]$ streng monoton fällt. Dann gilt aber wegen (5.5) und (5.8), der Stetigkeit der $U_\ell(x)$ und schließlich (5.6), daß es Punkte y_r, $r=1(1)m$, gibt mit

$$(5.9) \qquad -1 < x_1^{(m)} < y_1 < x_1^{(n)} < x_2^{(m)} < y_2 < x_2^{(n)} < \ldots < x_n^{(n)} < x_m^{(m)} < y_m < 1$$
$$\text{sodaß } U_m(y_r) = U_n(y_r), \quad r=1(1)m .$$

Da es somit m verschiedene Schnittpunkte y_r, $-1 < y_r < +1$, $(r=1(1)m)$ gibt, wechselt das Differenzpolynom $U_{n+1}(x) - U_n(x)$, $n \in \mathbb{N}$, im Innern von $[-1,+1]$ sein Vorzeichen $n+1$ mal, und unser Kriterium Satz 4.1 liefert die Behauptung. Satz 4.2 garantiert sogar, daß die Rivlin-Funktion aus $C^\infty[a,b]$ gewählt werden kann. Für $n=0$ folgt die Behauptung aufgrund des Kriteriums Satz 4.1 unmittelbar aus (5.6) und (5.7) wegen $U_o(x) \equiv 1$.

-- q.e.d. --

Im Anschluß an die Behauptung 1 und 2 stellt sich natürlich sofort die Frage, inwieweit für bestimmte Werte $n \in \mathbb{N}_o$ auch gilt, daß $L(U_n, U_{n+1}, U_{n+2})$ ungleich der leeren Menge ist. Es stellt sich also das Problem, ob die notwendigen Kriterien (NB) - Satz 3.2 - , (NL) - Satz 3.4.c - und (RB) - Satz 3.5 - für Tripel aufeinanderfolgender Tschebyscheff-Polynome 2. Art erfüllt sind.

Behauptung 3: Jedes Tripel (U_n, U_{n+1}, U_{n+2}), $n \geq 0$, erfüllt die Nullstellenbedingung (NB).

Diese Aussage ist eine unmittelbare Folge von Behauptung 1 und 2.

Behauptung 4: Jedes Tripel (U_n, U_{n+1}, U_{n+2}), $n \geq 0$, erfüllt die Randbedingung (RB), und zwar mit $R_1 = +1$.

Beweis: Gemäß (5.6) gilt

$$U_{k+\ell}(1) - U_\ell(1) = k+\ell+1-(\ell+1) = k, \quad k>0, \ell \geq 0,$$

und somit

$$\operatorname{sgn}\{U_{n+2}(1)-U_n(1)\} = \operatorname{sgn}\{U_{n+2}(1)-U_{n+1}(1)\} = \operatorname{sgn}\{U_{n+1}(1)-U_n(1)\} = +1.$$

Ferner ist mit (5.7)

$$U_{n+2}(-1)-U_{n+1}(-1) = \begin{cases} 2n+5 & > 0 \text{ falls } n \text{ gerade} \\ -(2n+5) & < 0 \text{ falls } n \text{ ungerade} \end{cases},$$

entsprechend ist

$$U_{n+1}(-1)-U_n(-1) = \begin{cases} -(2n+3) & < 0 \text{ falls } n \text{ gerade} \\ 2n+3 & > 0 \text{ falls } n \text{ ungerade} \end{cases}.$$

Somit gilt

$$\operatorname{sgn}\{U_{n+2}(-1)-U_{n+1}(-1)\} = -\operatorname{sgn}\{U_{n+1}(-1)-U_n(-1)\}, \quad n \geq 0.$$

Behauptung 5: Jedes Tripel (U_n, U_{n+1}, U_{n+2}), $n \geq 0$, erfüllt die Bedingung (NL) über die relative Nullstellenlage, genauer gilt

(5.10) $\quad x_1 < \min(x_2, y_1)$ und $\max(x_{i+1}, y_i) < \min(x_{i+2}, y_{i+1})$, für $1 \leq i \leq n$.

Dabei sind $x_i (i=1(1)n+2)$ bzw. $y_i (i=1(1)n+1)$ die der Größe nach wachsend angeordneten Nullstellen der Differenzpolynome $U_{n+2}(x)-U_n(x)$ resp. $U_{n+1}(x)-U_n(x)$.

Beweis: Daß die Differenzen $U_{n+2}(x)-U_n(x)$ bzw. $U_{n+1}(x)-U_n(x)$ jeweils lauter einfache Nullstellen haben, wurde in den Beweisen zu Beh. 1u.2 gezeigt. Beachtet man (5.5), so ist (5.10) gleichwertig mit der Bedingungsgruppe

$$(5.10a) \quad x_1 < y_1$$
$$(5.10b) \quad x_{i+1} < y_{i+1}, \quad i=1(1)n$$
$$(5.10c) \quad y_i < x_{i+2}, \quad i=1(1)n .$$

Gemäß (5.2) hat $U_{n+2}(x)-U_n(x)$ die Nullstellen

$$x_i = -\cos[(2i-1)\pi / (2n+4)], \quad i=1(1)n+2,$$

und gemäß (5.9) gilt die Abschätzung

$$(5.11) \quad x_i^{(n+1)} = -\cos[i\pi/(n+2)] < y_i < -\cos[i\pi/(n+1)] = x_i^{(n)}.$$

Wegen $1/(2n+4)<1/(2+n)$ ist somit $x_1<x_1^{(n+1)}$, also $x_1<y_1$, d.h. es gilt (5.10a). Wegen

$$[2(i+1)-1]/[2(n+2)] < (i+1)/(n+2)$$

folgt aufgrund (5.11) unmittelbar $x_{i+1}<y_{i+1}$, also (5.10b). Entsprechend ergibt sich wegen (5.11) aus

$$i/(n+1) < [2(i+2)-1]/(2n+4), \text{ wo } 1 \leq i \leq n \text{ ist},$$

sofort $y_i<x_{i+2}$, also (5.10c). Insgesamt also gilt die Behauptung.

-- q.e.d. --

Die Behauptungen 3-5 zeigten, daß für jedes Tripel (U_n, U_{n+1}, U_{n+2}), $n \in \mathbb{N}_o$, gewisse notwendige Kriterien für die Lösbarkeit des Rivlin-Problems erfüllt sind. Abschließend soll nun noch für zwei spezielle Tripel von Tschebyscheff-Polynomen 2. Art die Existenz einer Rivlin-Funktion nachgewiesen werden. Zunächst wird der Charakterisierungssatz 4.4 angewendet:

Behauptung 6: Für das spezielle Tripel (U_o^*,U_1^*,U_2^*) der geschifteten T-Polynome 2. Art existiert eine Rivlin-Funktion bezüglich des Intervalls $[0,1]$.
Dabei ist definiert $U_n^*(x) := U_n(2x-1)$, $n \geq 0$.

Beweis: Wir zeigen, daß die charakterisierenden Bedingungen des Satzes 4.4 erfüllt sind. - Es ist

$$U_o^*(x) = 1,\ U_1^*(x) = 4x-2,\ U_2^*(x) = 16x^2-16x+3.$$

Elementare Rechnung ergibt, daß gilt

$$U_1^*(z_i) = U_2^*(z_i),\ (i=1,2),\ \text{für } z_1=(5-\sqrt{5})/8 \text{ und } z_2=(5+\sqrt{5})/8.$$

Offensichtlich gilt $0 < z_1 < z_2 < 1$ und Bedingung Satz 4.4(a) ist erfüllt. Wegen $c_2 > 0$ und $b_1 > 0$ muß in Satz 4.4(b) erfüllt sein:

$$(z_1+z_2)/2 < (a_o-b_o)/b_1 < (1-z_2)/2.$$

Dies gilt in der Tat, denn $5/8 < 3/4 < (13+\sqrt{5})/16$.

-- q.e.d. --

Als Letztes wird nun das spezielle Tripel (U_1,U_2,U_3) von Tschebyscheff-Polynomen zweiter Art daraufhin untersucht, ob hierfür im Definitionsintervall $[-1,+1]$ eine Rivlin-Funktion existiert. Dieses Tripel ist zugleich ein Beispiel für die Anwendbarkeit unseres Satzes 4.3 in einem Fall, in dem die Anwendung des Satzes von D.SPRECHER (Satz 4.4) nicht möglich ist.

Behauptung 7: Für das Polynomtripel (U_1,U_2,U_3) sind alle Bedingungen des Satzes 4.3 erfüllt, d.h. es gilt
$L(U_1,U_2,U_3) \neq \phi$.

Beweis: Es ist $d_{21}(x)=U_3(x)-U_2(x)=8x^3-4x^2-4x+1$
$d_{20}(x)=U_3(x)-U_1(x)=8x^3-6x$
$d_{10}(x)=U_2(x)-U_1(x)=4x^2-2x-1$

Den Verlauf der Differenzpolynome zeigt BILD 5.
Die Funktionen $\{1,x,x^2,x^3\}$ bilden ein erweitertes Haarsystem; und es ist $U_1 \in \text{span}(1,x)$, $U_2 \in \text{span}(1,x,x^2)$, $U_3 \in \text{span}(1,x,x^2,x^3)$, wobei jeweils die höchsten Polynomkoeffizienten ungleich Null sind.
Satz 4.3 ist also mit r=1 anwendbar. Alle im Folgenden verwendeten Bezeichnungen entsprechen denen von Satz 4.3. ——
(Die numerischen Berechnungen erfolgten auf der Rechenanlage TR 4 des Leibniz-Rechenzentrums München. Die Zahlenangaben haben einen abs. Fehler $\leq 5 \cdot 10^{-7}$.)

1.: Bedingungsgruppe 1 von Satz 4.3

Daß diese Bedingungen erfüllt sind, folgt als Spezialfall mit n=1 aus den Behauptungen 3-5. Im Einzelnen gilt:

zu (NB): $d_{21}(x)$ hat die Nullstellen z_1=-0.623 490
z_2= 0.222 521
z_3= 0.900 969

$d_{20}(x)$ hat die Nullstellen x_1=-0.866 025 $=-\sqrt{3}/2$
x_2= 0.000 000
x_3= 0.866 025 $=+\sqrt{3}/2$

$d_{10}(x)$ hat die Nullstellen y_1=-0.309 017 $=(1-\sqrt{5})/4$
y_2= 0.809 017 $=(1+\sqrt{5})/4$

zu (NL): Es ist $-1=a<x_1<z_1<y_1<x_2<z_2<y_2<x_3<z_3<b = +1$

zu (RB): Es ist $d_{21}(-1)= -7$, $d_{20}(-1)= -2$, $d_{10}(-1)= +5$
$d_{21}(+1)= +1$, $d_{20}(+1)= +2$, $d_{10}(+1)= +1$
Somit ist in (RB): $R_1 = +1$ und $R_2 = -1$

2.: Bedingungsgruppe 2 von Satz 4.3

(A): Die Extrema der Differenzpolynome im Innern des Intervalls [-1,+1] liegen wie folgt:

bei $d_{21}(x)$: z_{max}=-0.274 292 , $d_{21}(z_{max})$ = 1.631 130
z_{min}=+0.607 625 , $d_{21}(z_{min})$ =-1.112 612
$d_{21}(x)$ ist streng monoton steigend in $[a,z_{max}]$, $[z_{min},b]$ und streng monoton fallend in $[z_{max},z_{min}]$.

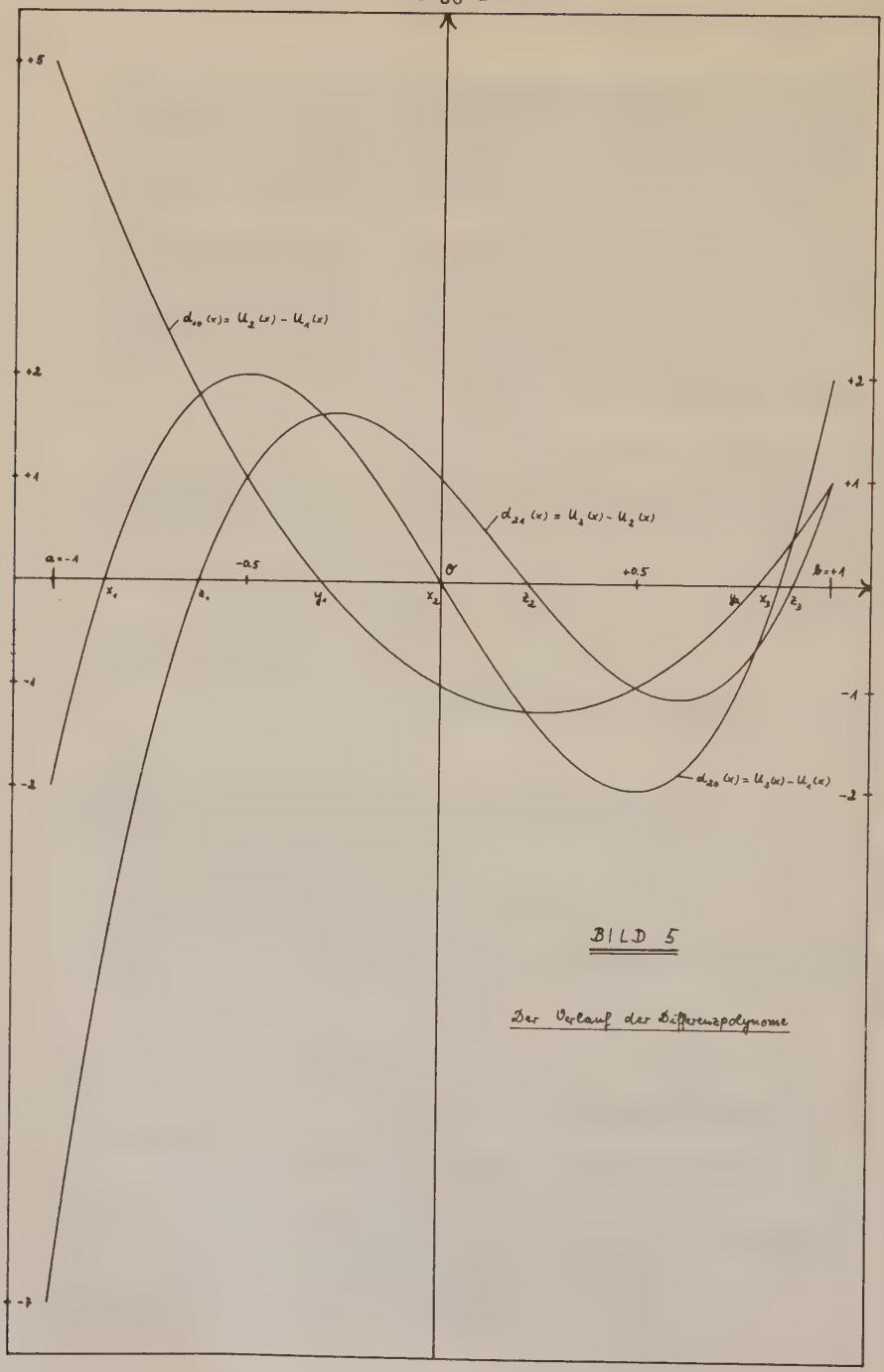

BILD 5

Der Verlauf der Differenzpolynome

bei $d_{20}(x)$: x_{max}= -0.500 000 , $d_{20}(x_{max})$ = +2.000 000
x_{min}= +0.500 000 , $d_{20}(x_{min})$ = -2.000 000
$d_{20}(x)$ ist streng monoton steigend in
$[a,x_{max}],[x_{min},b]$
und streng monoton fallend in
$[x_{max},x_{min}]$.

bei $d_{10}(x)$: y_{min}= +0.250 000 , $d_{10}(y_{min})$ = -0.125 000
$d_{10}(x)$ ist streng monoton steigend in $[y_{min},b]$
und streng monoton fallend in $[a,y_{min}]$.

(B): Zur Bedingung 2a von Satz 4.3

Es ist $\overline{D}_1 = (x_1,y_1)$, $\underline{D}_2 = (x_2,y_2)$, $\overline{D}_3 = (x_3,b]$

ferner $J_1^{21} = [a,z_1)$, $\overline{J}_2^{21} = (z_1,z_2)$, $J_3^{21} = (z_2,z_3)$, $\overline{J}_4^{21} = (z_3,b]$.

Beachtet man die Lage der Extrema (vergl. (A)) und die Werte der Differenzpolynome dort und auch an den Intervallenden von [-1,+1], so ergibt sich sofort:

$$n_1 = d_{10}(1)=1,\ n_2=d_{20}(1)=2,\ n_3=d_{21}(1)=1$$

Mit der folgenden Wahl der Konstanten k_i, i=1(1)3, ist dann die Bedingung 2a erfüllt:

$$k_1=0.674\ 286,\ k_2=0.965\ 714,\ k_3=0.291\ 428$$

(C): Zu den Bedingungen 2b von Satz 4.3

Folgende Wahl von U erfüllt die Bedingungen:

$U = [a_1,b_1]$ mit a_1=0.862 679 und b_1=0.933 853,

denn:

(a) Die Größen a_1,b_1 sind gerade so bestimmt, daß gilt $-d_{21}(a_1)=k_3=-d_{21}(b_1)$. Wegen der Monotonie von $d_{21}(x)$ in U (vergl.(A)) gilt somit
$|d_{21}(x)| \leq k_3$ für alle $x \in U$.

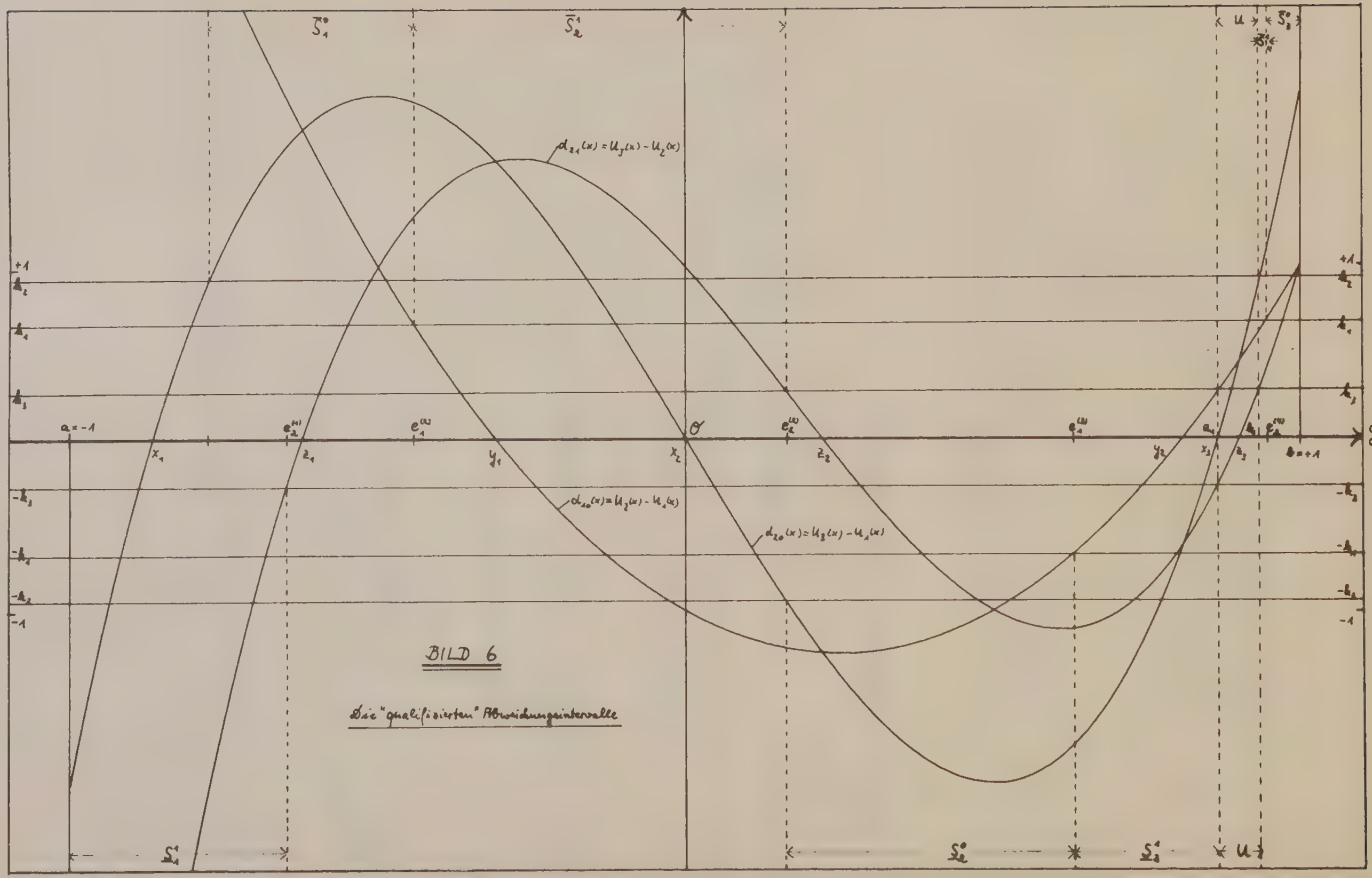

BILD 6

BILD 6

Die "qualifizierten" Abweichungsintervalle

(b) Es ist $d_{20}(a_1)=-0.039\ 928>-k_2$ und $d_{20}(b_1)=0.912\ 048<k_2$. Ebenfalls (vergl. (A)) wegen der Monotonie von $d_{20}(x)$ in U gilt $|d_{20}(x)|\leq k_2$ für alle $x\in U$.

(D): Zu den Bedingungen 2c von Satz 4.3

Es werden Intervalle $E_i=[e_1^{(i)},e_2^{(i)}]$, i=1(1)4, wie folgt definiert

$$e_1^{(1)} = a = -1 \quad , \quad e_2^{(1)} = -0.650\ 390$$
$$e_1^{(2)} = -0.443\ 593 \ , \quad e_2^{(2)} = +0.159\ 773$$
$$e_1^{(3)} = +0.629\ 379 \ , \quad e_2^{(3)} = a_1$$
$$e_1^{(4)} = b_1 \quad , \quad e_2^{(4)} = b = +1$$

und es wird gezeigt, daß diese Intervalle den Bedingungen 2c von Satz 4.3 genügen:

(a) Es gilt: $\underline{E}_1^{21}= E_1$, $\overline{E}_2^{21}= E_2$, $\underline{E}_3^{21}= E_3$ und $\overline{E}_4^{21}= E_4$, denn es ist:

$d_{21}(e_2^{(1)})=d_{21}(e_2^{(3)})=-k_3$ und $e_2^{(1)}<z_1, z_{min}<e_2^{(3)}<z_3$,

ferner

$d_{21}(e_2^{(2)})=d_{21}(e_1^{(4)})=k_3$ und $z_{max}<e_2^{(2)}<z_2, e_1^{(4)}>z_3$,

weiterhin

$d_{21}(e_1^{(2)})=1.288\ 970>k_3$ und $z_1<e_1^{(2)}<z_{max}$,

sowie schließlich

$d_{21}(e_1^{(3)})=-1.107\ 521<-k_3$ und $z_{min}<e_1^{(3)}<e_2^{(3)}=a_1$.

(b) Es gilt $\{x\in\overline{E}_i^{21}$ mit $d_{10}(x)\leq k_1\} = E_i$ für i=2 und i=4, sowie

$\{x\in\underline{E}_i^{21}$ mit $d_{10}(x)\geq -k_1\} = E_i$ für i=1 und i=3,

denn es ist

$$d_{10}(e_1^{(2)}) = -d_{10}(e_2^{(3)}) = d_{10}(e_2^{(4)}) = k_1$$

und $$-1 < e_1^{(1)} < e_2^{(1)} < e_1^{(2)} < y_1 < e_2^{(2)} < y_{min} < e_1^{(3)}.$$

3. Die "qualifizierten" Abweichungsintervalle

Die vorstehenden Punkte 1. und 2. schließen den Beweis der Behauptung 7 ab.
Hier sollen nun am vorliegenden Beispiel noch die "qualifizierten" Abweichungsintervalle angegeben werden, das sind die Intervalle, in denen (vergleiche die Bemerkungen zu Beginn des Beweisteils der Hinlänglichkeit in Satz 4.3) die Extremalpunkte (→ k-Abweichungsintervalle) der zu konstruierenden k-Nullalternierenden f_k, k=1(1)3, gewählt werden können. Siehe hierzu auch das BILD 6.

(A): qualifizierte Abweichungsintervalle für f_1:

$\bar{S}_1^1 := [c_1, e_1^{(2)}]$, wo $c_1 = -0.770\,246$, denn:

(a) $\operatorname{sgn} d_{10}(x) = \operatorname{sgn} d_{20}(x) = +1$ für $x \in \bar{S}_1^1$

(b) $d_{20}(c_1) = k_2$, $d_{20}(e_1^{(2)}) = 1.963\,255 > k_2$ und

$x_1 < c_1 < x_{max} < e_1^{(2)} < x_2$,

somit $|d_{20}(x)| \geq k_2$ für alle $x \in \bar{S}_1^1$

(c) $d_{10}(e_1^{(2)}) = k_1$ und $-1 < e_1^{(2)} < y_1$, somit $|d_{10}(x)| \geq k_1$ für $x \in \bar{S}_1^1$

$\underline{S}_2^1 := [c_2, e_1^{(3)}]$, wo $c_2 = +0.166\,667$, denn:

(a) $\operatorname{sgn} d_{10}(x) = \operatorname{sgn} d_{20}(x) = -1$ für $x \in \underline{S}_2^1$

(b) $d_{10}(c_2) = -0.674\,286 < -k_1$, $d_{10}(e_1^{(3)}) = -k_1$ und

$y_1 < c_2 < y_{min} < e_1^{(3)} < y_2$,

somit $|d_{10}(x)| \geq k_1$ für alle $x \in \underline{S}_2^1$

(c) $d_{20}(c_2) = -k_2$, $d_{20}(e_1^{(3)}) = -1.781\,807 < -k_2$ und

$x_2 < c_2 < x_{min} < e_1^{(3)} < x_3$,

somit $|d_{20}(x)| \geq k_2$ für alle $x \in \underline{S}_2^1$

$\overline{S}_3^1 := [e_2^{(4)}, b]$, denn:

(a) $\operatorname{sgn} d_{10}(x) = \operatorname{sgn} d_{20}(x) = +1$ für $x \in \overline{S}_3^1$

(b) $d_{10}(e_2^{(4)}) = k_1$ und $y_2 < e_2^{(4)} < b = +1$, somit $|d_{10}(x)| \geq k_1$ für $x \in \overline{S}_3^1$

(c) $d_{20}(e_2^{(4)}) = 1.059\ 602 > k_2$ und $x_3 < e_2^{(4)} < b$, somit $|d_{20}(x)| \geq k_2$ für $x \in \overline{S}_3^1$

(B): qualifizierte Abweichungsintervalle für f_2:

Gemäß Beweisteil 2.(D) gilt:

$$\underline{S}_1^2 = E_1,\ \overline{S}_2^2 = E_2,\ \underline{S}_3^2 = E_3,\ \overline{S}_4^2 = E_4$$

(C): qualifizierte Abweichungsintervalle für f_3:

In $U = [a_1, b_1]$, (vergl. Beweisteil 2.(C)), wähle man fünf verschiedene Punkte $a_1 < h_1 < h_2 < h_3 < h_4 < h_5 < b_1$, etwa $h_i := 0.873\ 000 + (i-1) \cdot 0.010\ 000$, $i=1(1)5$. Wegen der Wahl von U kann gesetzt werden $\underline{S}_i^3 = h_i$ für $i \in \{1,3,5\}$ und $\overline{S}_i^3 = h_i$ für $i \in \{2,4\}$.

LITERATUR

ABRAMOWITZ,M.; STEGUN,I. [1968] Handbook of Mathematical Functions. National Bureau of Standards, Appl. Math.Series 55,7th printing,May 1968, Washington

BERNSTEIN,S.N. [1912] Sur l'ordre de la meilleure approximation des fonctions continues par les polŷnomes de degré donné. Mémoire l'Acad.Royale de Belgique 4, S. 1-104

BERNSTEIN,S.N. [1938] Sur le problème inverse de la théorie de la meilleure approximation de fonctions continues. Comptes Rendus Acad.Sci. Paris 206, S. 1520-1523

BROSOWSKI,B. [1968] Nichtlineare Tschebyscheff-Approximation. B.I.-Hochschulskripten, Bd. 808/808a. Bibliographisches Institut Mannheim

BROSOWSKI,B. [1969] Nichtlineare Approximation in normierten Vektorräumen. Erschienen in "Abstrakte Räume und Approximation", hrsg. von Butzer-Nagy, ISNM Vol.10 (1969), S.140-162. Birkhäuser Verlag, Basel-Stuttgart.

BROSOWSKI,B.; WEGMANN,R. [1970] Charakterisierung bester Approximationen in normierten Vektorräumen. J.of Approx. Theory 3, S. 369-397

BUTZER,P.L.; SCHERER,K. [1969]
Über die Fundamentalsätze der klassischen Approximationstheorie in abstrakten Räumen.
Erschienen in "Abstrakte Räume und Approximation", hrsg von Butzer-Nagy, ISNM Vol. 10 (1969), S.113-125.
Birkhäuser Verlag, Basel-Stuttgart

DE RHAM,G. [1955]
Varietés differentiables.
Hermann u. Co., Paris 1955

DEUTSCH,F.; MORRIS,P.; SINGER,I. [1969]
On a problem of T.J. Rivlin in Approximation Theory.
J. of Approx. Theory 2, S. 342-352

HOFFMANN,K.-H. [1968]
Über nichtlineare Tschebyscheff-Approximation mit Nebenbedingungen.
Dissertation, Universität München, 1968

JACKSON,D. [1930]
The Theory of Approximation.
AMS Colloquium Publications, Vol. XI, New York 1930

KARLIN,S. [1968]
Total positivity, Vol. I
Stanford University Press, California

KARLIN,S.; STUDDEN,W.J. [1966]
Tchebycheff Systems: With Applications in Analysis and Statistics.
Intersciense Publishers, New York-London-Sydney

KÖTHE,G. [1966]
Topologische lineare Räume I, 2.Auflage.
Springer Verlag, Berlin-Heidelberg-New York

LJUSTERNIK,L.A.; SOBOLEW,W.I. [1965]
Elemente der Funktionalanalysis.
Akademie-Verlag, Berlin

NITSCHE,J. [1968] Bemerkungen zur Approximationsgüte bei projektiven Verfahren. Math.Zeitschrift 106, S.327-331

RIVLIN,T.J. [1969] New and unsolved problems, No.14: Best algebraic approximation. Erschienen in "Abstrakte Räume und Approximation", hrsg. von Butzer-Nagy, ISNM Vol. 10 (1969), S. 421 Birkhäuser Verlag, Basel-Stuttgart.

SINGER,I. [1969] Remark on a paper of Y.Ikebe. Proceedings of the AMS (21), S.24-26

SINGER,I. [1970] Best approximations in Normed Linear Spaces by Elements of Linear Subspaces. Springer Verlag, Berlin-Heidelberg-New York

SPRECHER,D. [1969] Simultaneous Best Approximations with Two Polynomials. J. of Approx. Theory 2, S.384-388

SPRECHER,D. [1969a] On Simultaneous Chebyshev Approximations with Polynomials. (Unveröffentlichtes Manuskript)

STEČKIN,S.B. [1951] On the order of best approximations of continuous functions. (Russisch). Izv. Akad. Nauk SSSR 15, S. 219-242

STOER,J. [1967] Über die Existenz von linearen Approximationsoperatoren. Erschienen in "Funktionalanalysis, Approximationstheorie, numerische Mathematik", ISNM Vol. 7(1967), S. 55-57. Birkhäuser Verlag, Basel-Stuttgart

van der WAERDEN,B.L. [1964] Algebra, 1. Teil
Springer Verlag, Berlin-Göttingen Heidelberg

ZAMANSKY,M. [1949] Classes de Saturation de certains procédés d'approximation de séries de Fourier de fonctions continues et applications à quelques problèmes d'approximation.
Ann. sci. École Normale Sup. 66, S. 19-93

ZYGMUND,A. [1945] Smooth functions.
Duke Math. J. 12, S. 47-76

LEBENSLAUF

Am 29. Mai 1943 wurde ich in Recklinghausen geboren. Meine Eltern sind der Tischlermeister Heinrich Hegering und seine Ehefrau Maria geb. Schneider.

Nach vier Jahren Volksschule besuchte ich den Neusprachlichen Zweig des Hittorf-Gymnasiums in Recklinghausen, wo ich am 27. 2. 1962 die Reifeprüfung bestand. Anschließend studierte ich vom Sommersemester 1962 bis zum Wintersemester 1965/66 in Münster und vom Sommersemester 1966 bis zum Sommersemester 1968 an der Universität München das Fach Mathematik. Am 30. 7. 1968 legte ich die Diplom-Hauptprüfung in Mathematik (wirtschaftswiss. Richtung) ab. Meine Diplomarbeit schrieb ich bei Herrn Professor Dr. G. Hämmerlin.

Seit November 1968 bin ich als wiss. Angestellter am Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften beschäftigt.

Von meinen akademischen Lehrern hat insbesondere Herr Professor Dr. G. Hämmerlin meinen beruflichen Werdegang entscheidend beeinflußt. Dafür sage ich ihm herzlichen Dank.